R E P O R T   R E S U M E S

ED 015 678   56   EM 005 973

SCRAMBLED VERSUS ORDERED SEQUENCING IN COMPUTER-ASSISTED INSTRUCTION.

BY- WODTKE, KENNETH H. AND OTHERS
PENNSYLVANIA STATE UNIV., UNIVERSITY PARK

REPORT NUMBER PN-1435   PUB DATE JUL 67
REPORT NUMBER BR-5-8334
GRANT OEG-1-6-058334-1819
EDRS PRICE MF-$0.50 HC-$4.52   111P.

DESCRIPTORS- *COMPUTER ASSISTED INSTRUCTION, *ANALYSIS OF VARIANCE, *ABILITY, *TIME FACTORS (LEARNING), *SEQUENTIAL APPROACH

THE INVESTIGATORS HYPOTHESIZED THAT SCRAMBLING AN INSTRUCTIONAL SEQUENCE BASED ON HIERARCHICAL CONCEPTS WOULD BE HIGHLY DETRIMENTAL TO LEARNING, BUT THAT SCRAMBLING A SEQUENCE OF RELATIVELY UNRELATED FACTS WOULD NOT AFFECT LEARNING. IT WAS ALSO EXPECTED THAT SCRAMBLING WOULD BE MOST DETRIMENTAL TO LOW ABILITY STUDENTS, BUT WOULD NOT AFFECT HIGH ABILITY STUDENTS. A PILOT STUDY AND TWO EXPERIMENTS TESTED THE EFFECTS OF SCRAMBLING VERSUS ORDERED SEQUENCES IN PROGRAMS TEACHING MODERN MATHEMATICS (DEPENDENT ON A HIERARCHY OF CONCEPTS) AND EAR ANATOMY TO COLLEGE STUDENTS BY COMPUTER-ASSISTED INSTRUCTION. ANALYSIS OF VARIANCE SHOWED NON-SIGNIFICANT EFFECTS OF A SCRAMBLED ANATOMY SEQUENCE, BUT DETRIMENTAL EFFECTS OF A SCRAMBLED MATH SEQUENCE WERE MUCH SMALLER THAN EXPECTED. ALTHOUGH THE SCRAMBLED MATH SEQUENCE INCREASED ERRORS AND INSTRUCTIONAL TIME, ACHIEVEMENT LEVEL, AS COMPARED WITH AN ORDERED SEQUENCE GROUP ON AN IMMEDIATE POST TEST, WAS NOT AFFECTED. ALSO, TENTATIVE SUPPORT WAS GIVEN TO THE HYPOTHESIS THAT AN ORDERED SEQUENCE IS MORE IMPORTANT TO THE LEARNING OF LOW APTITUDE STUDENTS. (LH)

EM 005973 A
5-8334
56

C·A·I

COMPUTER ASSISTED INSTRUCTION LABORATORY

COLLEGE OF EDUCATION · CHAMBERS BUILDING

THE PENNSYLVANIA STATE UNIVERSITY · UNIVERSITY PARK, PA.

SCRAMBLED VERSUS ORDERED SEQUENCING IN COMPUTER-ASSISTED INSTRUCTION

Kenneth H. Wodtke, Bobby R. Brown,
Harold R. Sands and Patricia Fredericks

July, 1967

SCRAMBLED VERSUS ORDERED SEQUENCING IN

COMPUTER-ASSISTED INSTRUCTION

Project No. 1435
Grant No. OEG1-6-058334-1819

by

Kenneth H. Wodtke, Bobby R. Brown,
Harold R. Sands and Patricia Fredericks

July, 1967

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

The Pennsylvania State University
University Park, Pennsylvania

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

RANDOM VERSUS ORDERED SEQUENCING IN COMPUTER-ASSISTED INSTRUCTION

Introduction

The ordered or " gical" sequencing of learning materials has been most frequently mentioned as one of the primary advantages of programed instruction. Recently a number of papers have appeared which present inconclusive and frequently contradictory results concerning the advantages or disadvantages of ordered sequences of items in a program compared to random sequences of items. Some of these studies report superior learning for carefully organized instructional programs, while others report no differences between random and ordered presentation conditions. In the writer's judgment, this basic conflict must be resolved before research can proceed to the investigation of more subtle sequencing and organizational variables in programed instruction.

The present project examined the question of random versus ordered sequencing within the context of several other variables thought to be important in determining the effects of program organization. Few previous studies on this problem employed factorial designs to study the effects of important interactions between sequencing variables and other relevant variables. It is the investigators' general position that the benefits of careful sequences of learning materials will depend on the kinds of students being taught, the content of the instructional program, the kinds of educational outcomes being sought (for example, recall of facts as compared to the learning of principles and transfer to new situations), and the care and extent to which an experimental version of an instructional program is initially sequenced. A number of recent single-factor experiments have given the impression that carefully organized learning sequences in instructional programs "make no difference." The present study aimed at the more accurate statement that carefully organized learning sequences "make no difference," only <u>under certain conditions</u>, and that course organization is a highly critical variable in most educationally relevant situations. The relevant studies and rationale are developed more fully below.

The question of the importance of careful sequencing or conceptual organization of instructional programs is particularly important in the context of computer-assisted instruction which provides a high degree of flexibility in the planning, sequencing, and organization of course materials. To maximize the potential of computer-assisted instructional systems, unambiguous answers to questions concerning course sequencing must be found.

Related Literature and Rationale

There has been much recent exploratory research activity on the instructional potential of high speed electronic computers. Current projects at The Pennsylvania State University and Florida State University (1964) based on a computer-assisted instructional system developed by I.B.M. research workers at I.B.M.'s T. J. Watson Research Center are developing and testing college courses for presentation to students via a high speed electronic computer. This work stems from the earlier work of Uttal (1962) at I.B.M. Other approaches to computer-assisted instruction are exemplified by the work of Bitzer and Easley (1964), Braunfeld (1964), Stolurow and Davis (1963), and Suppes (1964). Although these different systems of computer-assisted instruction vary in some respects they all have in common a great flexibility of course organization and sequencing, and the capability of adapting course sequences to individual learners. Unfortunately, there is as yet little consistent evidence that even the simplest "carefully organized sequences" produce learning superior to randomized sequences. In spite of these inconsistencies of research results, it is hard to believe from an intuitive standpoint that randomized learning sequences produce learning equivalent to carefully organized learning sequences particularly for highly complex tasks.

B. F. Skinner (1958) and R. Glaser (1961) were among the first to elucidate the principle of small step, carefully sequenced items in instructional programs. Recently, proponents of branching programs have argued that a wide variety of appropriate instructional sequences may exist. In the typical branching program, learning sequences vary depending upon the student's own responses, thus, the branching program provides individualized instructional sequences. Recently, a number of studies have appeared in which random or scrambled item sequences in autoinstructional programs have been compared with ordered or "logical" item sequences. Roe, Case, and Roe (1962) compared the performance of 36 freshman psychology students on a program in elementary probability. Half of the students were assigned to a scrambled sequence, the other half to an "ordered" sequence. Nonsignificant differences were obtained in the criterion test performance of the two groups. The authors concluded: "The results of this small scale experiment, however, seem to indicate that college level students may not require the careful sequencing of autoinstructional items as had previously been supposed." (Roe, Case, & Roe, 1962, p. 104).

Levin and Baker (1963) compared the performance of two groups of second grade children in a scrambled and ordered program in informal geometry, and again nonsignificant differences were found. These authors concluded that "There was no evidence that item scrambling impaired learning." The writers later made an important point which has frequently been ignored in other studies. The posttest performance indicated that the "ordered" program failed to teach the material very effectively. That two poorly organized programs produce equivalent posttest achievement is not surprising. Studies of sequencing variables must provide some indication of the sequential characteristics of the "ordered" program.

Hamilton (1964) obtained a complex interaction between random versus ordered sequencing and specific (fill in the blank) versus nonspecific (read frames containing the correct answer) responding. These data indicated that the random-nonspecific combination was the most successful of all treatments. Unfortunately, however, these data are very difficult to interpret since the specific versus nonspecific response condition was confounded with the effects of information feedback. No information feedback was given in the specific response condition.

In contrast to results which question the importance of carefully sequenced instructional materials, are studies that have obtained significant differences associated with sequencing of material in complex verbal learning. Illustrative studies are those of Gagné (1962) and Evans (1960). Gagné presented a procedure for organizing tasks in terms of a knowledge hierarchy. According to this point of view complex tasks may be broken down into compoment prerequisite tasks which must be learned before successive steps in the task may be mastered. Gagné provided an illustration of his procedure for seven subjects on a number series task. The procedure involves working backwards from the requirements of the final task to prerequisite subtasks. This procedure provides a useful method for constructing and analyzing self-instructional programs. Evans (1960) compared programs constructed by the "Ruleg" system to programs constructed by less systematic methods. The Ruleg system produced the same degree of learning as the less systematic program, but less time was required to reach the obtained level.

The following comments can be made concerning the unresolved issue of sequencing in programed instruction:

1. An examination of the programs used in a number of studies reveals much variability in subject matter content. If a program teaches vocabulary or knowledge of terms and there are few interrelationships among the items then careful sequencing is not likely to be as important for such materials. Unrelated facts will probably be learned as well if presented in random order as in some specified sequence. On the other hand, performance on a program involving relationships between concepts, and the understanding and application of principles will be detrimentally

effected by random or scrambled presentation. To provide data on this hypothesis, the present study employed two programs which have been developed as part of Penn State's project in computer-assisted instruction (OE-4-16-010). These programs differed in the extent to which they taught principles and relationships as opposed to relatively unrelated facts.

2. The writers agree with Lumsdaine (1963) who has commented on the importance (in studying sequencing) of the susceptibility of stimulus materials to the utilization of verbal mediating responses. Some students undoubtedly provide their own conceptual organization to a scrambled program by linking varied parts of a randomized sequence of items with a verbal mediator. Since students vary considerably in their verbal mediational ability it is likely that an interaction exists between random versus ordered sequencing and level of student verbal ability. The effects of random sequencing are expected to be particularly detrimental to the learning of students exhibiting low verbal ability who would be unable to supply their own conceptual organization of the subject matter. A 2 x 2 factorial design was used to examine this interaction. The use of student verbal ability as one measured independent variable in the design also provided some indication of the sensitivity of the experiment as a whole since verbal ability is widely known to correlate with student achievement. The latter methodological suggestion has been made by Lumsdaine (1963). If one builds a "strong" variable into a research design and no significant effect is obtained for this variable, this would lead to the conclusion that the experiment was too insensitive to pick up differences due to other treatments.

3. The literature on sequencing of instructional programs also suggests that variations in the objectives of learning or student outcomes as measured by criterion tests is an important variable. Most of the criterion tests employed in previous studies (insofar as can be determined from the published reports) measured the student's recall of facts learned in the program. If this is all that is required, then scrambling or randomizing item sequence may not be expected to have a serious detrimental effect on learning. On the other hand, if one were to measure student achievement of a number of other learning objectives, such as, the understanding of a principle or relationship, or the student's ability to employ a principle in new problems, organized sequencing of program content might produce superior performance. In the present study two criterion tests were employed, one testing recall of factual material presented in the program, and the other testing the student's understanding of the principles involved and his ability to apply these principles to problems not contained in the original instructional program (e.g., a transfer task).

4. Finally, the investigators were convinced that the lack of a pretest was a methodological short coming in some earlier studies. The assumption has sometimes been made that students have "zero" knowledge about some content areas used in studies of programed instruction, and on this basis the pretest has been omitted. That the assumption of "zero" knowledge is highly untenable for adult or college populations has been demonstrated by Mager and Clark (1963). The present study employed a pre-posttest design to control the level of pretest performance.

Objectives

Although the objectives of the study have already been discussed to a considerable extent above, they may be briefly summarized as follows:

a. To determine under what conditions careful sequencing of instructional programs "make a difference" in student learning within the context of computer-assisted instruction. Following appropriate null hypothesis tests, it was expected that scrambled item sequencing would have a detrimental effect on the learning of program content in which the mastery of some concepts and principles were prerequisite to the mastery of other concepts, (e.g., a program containing a conceptual hierarchy).

b. To determine whether the effects of item sequencing depended in part upon the instructional outcomes desired. If the objective of instruction is recall of unrelated facts, item sequencing may be less important as a critical programing variable. When understanding of principles and transfer to new problems are desired, the effects of item sequencing may be more pronounced.

c. To determine whether random as compared to ordered item sequences have a differential effect on students of high as compared to low verbal ability. Randomized sequences were expected to have a more detrimental effect on the learning of low verbal ability students due to the inability of these students to reorganize the scrambled program on their own.

Pilot Investigation

This is the first in a series of investigations of the effects of course sequencing in CAI. The primary purpose of the pilot experiment was to investigate the interaction between student aptitude and scrambled versus ordered sequencing of instruction. In contrast to earlier investigations, the present study employed a fairly lengthy instructional program of considerable difficulty for the average college student. The material used involved the learning of principles, mathematical problem solving, and contained a large number of sequential dependencies among the concepts taught. The specific objectives and predictions of the experiment were as follows:

a) To determine under what conditions careful sequencing of instructional programs "make a difference" in student learning within the context of computer-assisted instruction. Following appropriate hypothesis tests, it was predicted that scrambled item sequencing would have a detrimental effect on student learning in a relatively lengthy, difficult program containing many sequential dependencies among concepts, e.g., when the mastery of some concepts and principles are prerequisite to the mastery of other concepts and principles.

b) To determine whether scrambled as compared to ordered item sequences have a differential effect on students of high- as compared to low-verbal aptitude. An aptitude by sequencing interaction effect was predicted. Scrambled item sequences were expected to have a more detrimental effect on the learning of low verbal ability students than on the learning of high verbal ability students. It was expected that students of low-verbal ability would not have the conceptual skills required to reorganize the scrambled material.

Methods and procedures

Description of instructional system (CAI) and course materials. The course used in the first experiment was a section of a modern mathematics course which was developed for CAI by the staff of the Computer Assisted Instruction Laboratory of the Pennsylvania State University.[1] The material selected contained instruction

---

[1]The writers would like to thank Professor Alan Riedesel and Marilyn Suydam of the Penn State Computer Assisted Instruction Laboratory who developed the original version of the Modern Mathematics program.

on the use of number systems with bases other than ten. This learning task offered the advantage of being relatively difficult for college students to learn, and the material was unfamiliar to most students. The ordered version of the program presented subsets of items in the following sequence: review of the base ten system; the concept of place value; the application of the concept of place value in base eight, base two, and base twelve number systems; transformations from one base to another; addition and subtraction in number systems with bases other than ten; and multiplication and division in number systems with bases other than ten. Previous experience with these course materials indicated that most undergraduate college students could complete instruction in approximately two and one-half to three hours with a mean error rate of about fifteen per cent.

The course materials used in the present study were prepared for CAI by means of a special computer language known as Coursewriter (Maher, 1964) developed by I.B.M. computer scientists at the Thomas J. Watson Research Center, Yorktown Heights, New York. Using the Coursewriter language, a course was programed including questions, problems, correct answers, incorrect answers, knowledge of results, and remedial branches, all of which were stored on magnetic discs to which the computer had selective access to any part with an access time of less than one second. The computer was programed to accumulate and store all student errors and response latencies, and these data were later retrieved for the investigators by means of a special program called Student Records. The scrambled sequence version of the number systems program was established by rearranging the sequence of frames according to a table of random numbers. The scrambled sequence was then entered and stored on the magnetic discs as a separate course.

The central computer used in the study was an I.B.M. 7010-1410 system located at I.B.M.'s Thomas J. Watson Research Center, Yorktown Heights, New York. The course materials in the form of questions, problems, prompts, etc., were teleprocessed over long distance telephone lines to student terminals on The Pennsylvania State University campus. The course was presented to students via an I.B.M. 1050 student terminal which consisted of a modified two-way electric typewriter, and a random access slide projector and tape recorder (the slide projector and tape recorder were not used in the present study). Questions and problems were typed out to the student, who typed his responses at the terminal. The student relayed his responses to the central computer which evaluated the response, provided knowledge of results, and sequenced the student to the next appropriate step in the course.

Subjects and procedures. Fifty-one undergraduate students in an introductory educational psychology class at The Pennsylvania

State University served as the Ss in the investigation. Ss with absolutely no previous typing experience were not included in the study. Two Ss were eliminated because a modern mathematics pretest indicated they had previous knowledge of number systems with bases other than ten. One other S was eliminated because his Scholastic Aptitude Test scores (SAT) were not available. These eliminations brought the total number of Ss to 48.

Subjects were then subdivided into high- and low-aptitude groups on the basis of their scores on the verbal Scholastic Aptitude Test (SAT). The mean of the high group was 612 and the mean of the low group was 435 (SAT employs standard scores based on a mean of 500 and a standard deviation of 100). The original plan of the pilot investigation was to assign Ss within each of the high and low aptitude groups at random to the scrambled or ordered instructional treatment conditions. Although approximately half of the Ss were assigned to treatments at random, the random assignment of a large number of Ss had to be altered due to a number of programing "bugs" which developed at the last minute in the scrambled sequence program. For this reason, a larger number of Ss who were scheduled for the early experimental sessions were run in the ordered sequence condition, and a larger number of Ss scheduled for the later experimental sessions were run in the scrambled sequence condition. The investigator carefully examined the two groups of subjects and in spite of the nonrandom assignment of some of the Ss, could find no evidence of selective factors which could account for the results obtained in the study. However, to provide additional assurance of the reliability of the results, a second independent replication of the study was conducted (see Experiment I).

The Ss reported to the Computer Assisted Instruction laboratory individually and were given Form A of a 23-item achievement test as a pre-treatment examination to test their prior knowledge of number systems with bases other than ten. Initially, each S was given a warm-up to familiarize him with the student typewriter terminal. After a warm-up period of about fifteen to thirty minutes, the S was allowed to begin instruction on the number systems program. At the completion of the program, S was given Form B of the 23-item achievement test on number systems. The reliability of Form B of the criterion measure estimated by the Hoyt technique was found to be .93 in an earlier study (Mitzel and Wodtke, 1965). The test-retest reliability of the criterion measure was also found to be .93 for a one-week interval between testings in the earlier investigation. Following the achievement posttest, all Ss completed a Student Reaction Inventory consisting of Semantic Differential type items (Osgood et. al., 1957) designed to measure the student's attitude towards CAI.

All Ss were allowed to complete the instructional materials at their own rate. Thirty-seven Ss were able to complete the program in one evening, while 11 had to return the following day to finish the material. Two Ss in the scrambled sequence group were scheduled to return the following day to complete instruction, but they failed to return. These two Ss seemed highly frustrated by the scrambled sequence program.

As previously mentioned, two Ss were eliminated because their pretest scores indicated prior knowledge of number systems. The remaining Ss who were included in the study achieved, on the average, one-half point on the pretest. Seventy per cent of the Ss obtained a score of zero on the pretest indicating that the students had little or no prior knowledge of number systems with bases other than ten.

The dependent variables of the study were achievement post-test scores, errors made in the program, total time taken to complete the program, mean response latency per frame, an efficiency score obtained by taking the ratio of criterion test performance to instructional time, and measures of the students' attitudes towards CAI. The data were analyzed by means of a two by two factorial analysis of variance design with unequal numbers of cases per subcell. One experimental factor consisted of high versus low aptitude; the other of scrambled versus ordered program sequence.

Results

A preliminary analysis indicated that although the high- and low-aptitude groups differed significantly on the verbal SAT measure, the scrambled and ordered sequence groups did not differ significantly in verbal ability as measured by the SAT. In addition, an analysis of Quantitative SAT scores produced nonsignificant differences among the four treatment groups employed in the study.

The distributions and the variances within groups of the dependent variables were examined to determine whether the assumptions underlying the analysis of variance had been met. None of the distributions appeared to deviate substantially from normality. Hartley's Maximum F-ratios were computed to test the assumption of homogeneity of variance. All of the F-ratios were nonsignificant except one. The F-ratio for the efficiency score was significant at less than the .01 level indicating the presence of heterogeneity of variance for this variable. In view of the results obtained by Boneau (1960) and Norton (1952) who found that heterogeneity of variance did not seriously bias either the t-test or F-ratio, the heterogeneity of variance for the efficiency score could not have seriously biased the results obtained in the present study.

Table 1 summarizes the results of the analyses of variance of three of the dependent variables, frequency of errors made in the program, per cent errors, and criterion test score. The results indicated that students in the scrambled sequence group made significantly more errors during instruction than the students in the ordered sequence group ($P < .001$). Since the students in the scrambled sequence group were more likely to encounter remedial segments of the program (due to their greater tendency to make errors) than the students in the ordered group, the scrambled sequence group actually responded to more questions than the ordered sequence group. The differences obtained in the total frequency of errors might have resulted from the fact that the students in the scrambled group simply responded to more questions and thus had more opportunity to make errors than the ordered group. To control for this possibility, an analysis was also computed based on per cent error scores. As shown in Table 1, this analysis also indicated that students in the scrambled sequence group made a significantly greater percentage of errors than the ordered sequence group. In spite of the highly significant sequencing main effect for frequency and percentage of errors, the sequencing main effect for the criterion test score was nonsignificant. Considered together, these results indicated that although the scrambled sequence students made significantly more errors during instruction than the ordered sequence Ss, they apparently improved their performance during instruction and, by the end of the course, they performed approximately at the same level as the ordered group on the criterion measure. A more detailed analysis of the frequency of errors made during instruction was undertaken in Experiment I to determine whether students in the scrambled sequence group showed improvement from the beginning to the end of the course.

The results reported in Table I also show an aptitude by sequencing interaction of borderline significance ($P < .10$) for the frequency of errors and criterion test scores. However, the interactions which were obtained for these variables did not result from a decrement in the performance of the low-aptitude group in the scrambled program as predicted, but from a decrement in the performance of the high aptitude Ss in the scrambled program. The results of the present study suggest the tentative conclusion that scrambling the instructional program had little or no effect on the performance of low-aptitude students, but produced a rather marked decrement in the performance of high-aptitude students. The graphs of the interactions for the frequency of errors and criterion test variables are shown in Figs. 1 and 2. Both of these figures show the sharp drop in performance of the high- aptitude students in the scrambled sequence program.

Table 1

Analyses of Variance of Frequency of Errors, Per Cent Errors, and Criterion Scores for High- and Low-Aptitude Students in Scrambled and Ordered Sequence Conditions

Source	d.f.	Frequency of Errors F-ratios	Per Cent Errors F-ratios	Criterion Score F-ratios
Aptitude	1	1.48	2.42	.27
Sequencing	1	12.65***	11.94**	1.40
Aptitude x Sequencing	1	3.96*	.69	3.62*
Error	44	(529.26)[a]	(801.60)[a]	(32.07)[a]

[a] Equals the mean square of the error term
* P is less than .10
** P is less than .01
*** P is less than .001

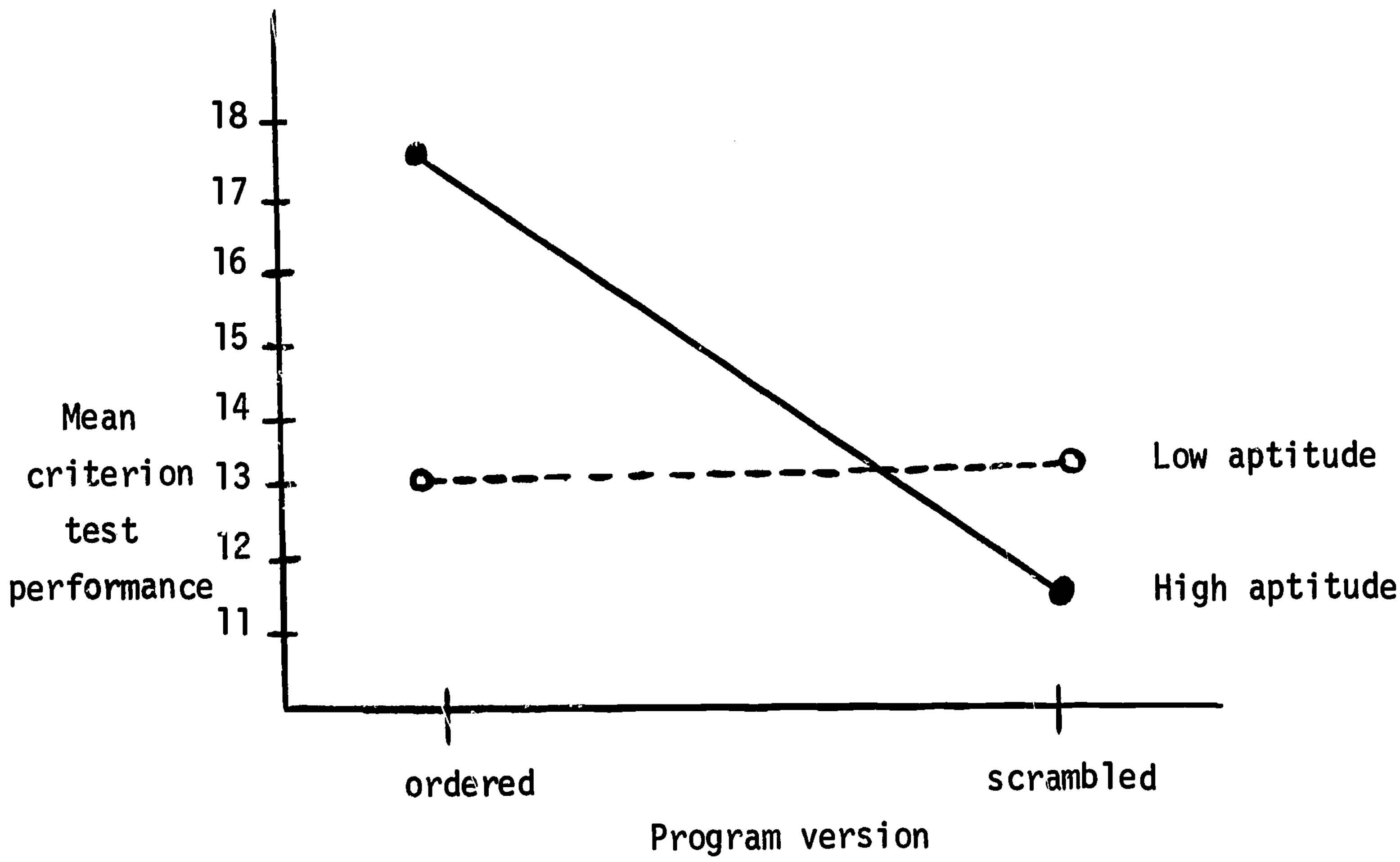


Fig. 1. Criterion test performance of high- and low-aptitude groups taught by scrambled and ordered instructional programs. (N's equaled: HA-ordered = 17, HA-scrambled = 9, LA-ordered = 14, LA-scrambled = 8)

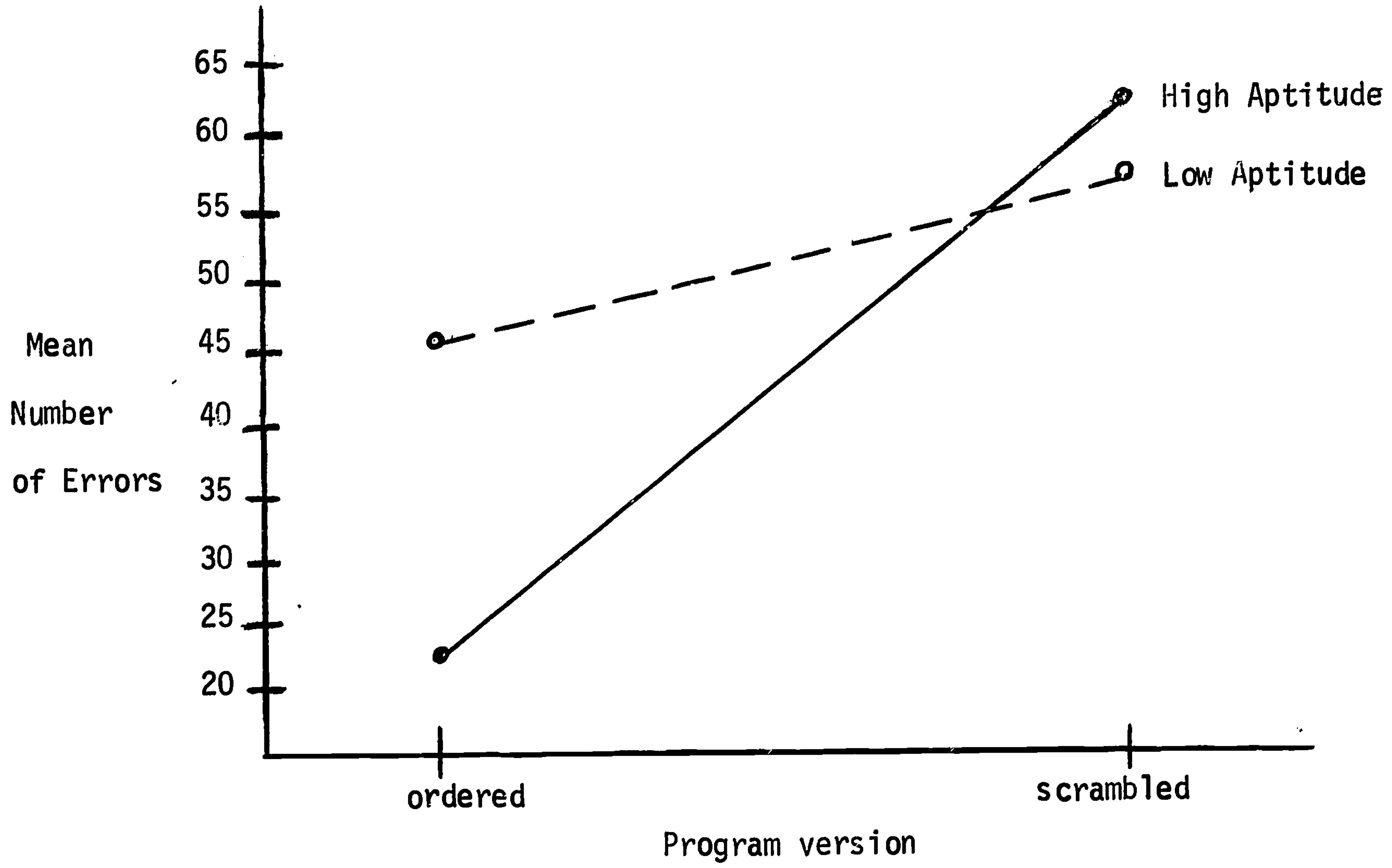


Fig. 2. Mean number of errors of high- and low-aptitude students taught ty scrambled and ordered programs. (N's equaled: HA-ordered = 17, HA-scrambled = 9, LA-ordered = 13, LA-scrambled = 9)

The analyses of the other dependent variables are consistent with the results already reported. Table 2 shows the analyses of variance summaries for total instructional time, average response latency per frame, d the efficiency score measures. Students in the scrambled sequence group took significantly more time to complete the program than students in the ordered sequence group ($P < .001$). The scrambled sequence group took on the average 45 minutes longer to complete the program than the ordered sequence group. The longer total instructional time taken by the scrambled sequence group was not simply the result of their responding to more questions than the ordered group. The results reported in Table 2 also indicated that the scrambled sequence group on the average took longer to respond to individual frames of the program. The sequencing main effect for the average latency variable was significant at less than the .01 level. The scrambled sequence students took on the average 2.05 minutes per frame to respond, while the students in the ordered sequence group averaged 1.59 minutes per response.

The efficiency score reflects the amount learned per unit of time as measured by the criterion measure of achievement. The efficiency score was obtained by taking the ratio of a student's criterion test performance to his total instructional time. The results reported in Table 2 show a sequencing main effect which was statistically significant at less than the .01 level. If one compares only the criterion test performance of the scrambled and ordered sequencing groups, as was reported in Table 1, it is interesting to note that the difference is non-significant. One possible explanation for this finding is that the additional instructional time and remedial frames taken by the students in the scrambled sequence group brought their performance up to a level comparable to that of the ordered sequence group. When the additional instructional time taken by the scrambled sequence group is taken into account by using an efficiency score, the scrambled sequence program is found

Table 2

Analyses of Variance of Total Instructional Time, Average Response Latency, and Efficiency Scores for High- and Low-Aptitude Students in Scrambled and Ordered Sequence Conditions

Source	d.f.	Total Time F-ratios	Ave. Resp. Latency F-ratios	Efficiency Score F-ratios
Aptitude	1	.79	1.63	2.07
Sequencing	1	14.33***	10.77**	7.50**
Aptitude x Sequencing	1	2.52	.02	4.75*
Error	44	(1591.03)[a]	(.22)[a]	(3475.43)[a]

[a] Equals the mean square of the error term
* P is less than .05
** P is less than .01
*** P is less than .001

to be less efficient than the ordered program, even though the two groups eventually reach approximately comparable levels of criterion test performance.

A statistically significant aptitude by sequencing interaction was again obtained in the analysis of the efficiency scores ($P < .05$). This result is consistent with the interactions reported for the error and criterion variables. The scrambled instructional program produced the largest decrement in the efficiency of performance of the high-aptitude students but this effect was due in part to the fact that the high-aptitude scrambled sequence Ss took the longest time to complete instruction. This interaction is shown graphically in Fig. 3, which shows the relatively sharp drop in the efficiency scores of the high-aptitude group in the scrambled sequence condition.

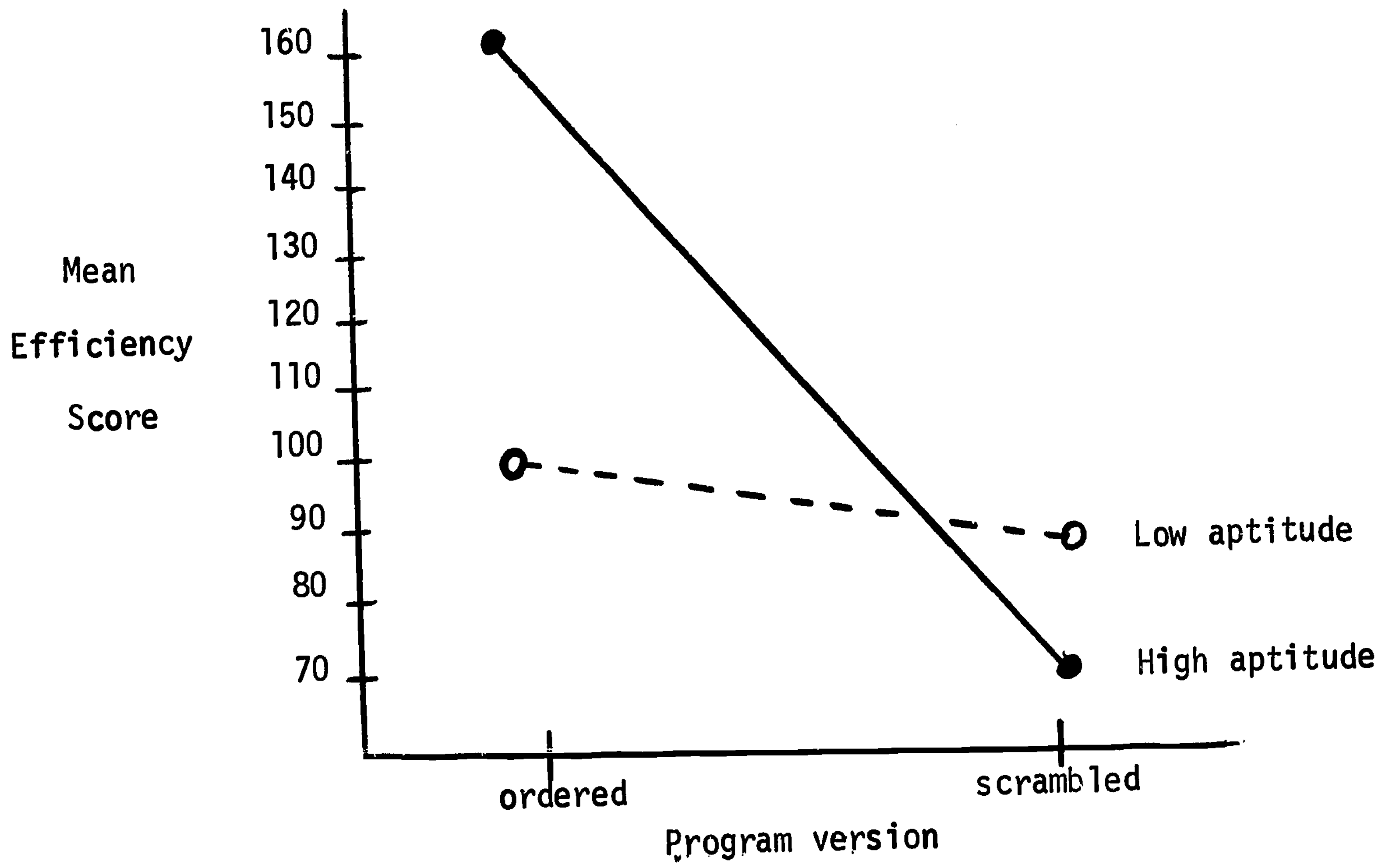


Fig. 3. Mean efficiency scores of High- and low-aptitude students taught by scrambled and ordered programs. (N's equaled: HA-ordered = 17, HA-scrambled = 9, LA-ordered = 13, LA-scrambled = 8)

In considering the three interactions depicted in Figs. 1 through 3, the reader should keep in mind that, in each case, the high-aptitude students in the scrambled sequence condition took the most instructional time but showed the poorest performance of all the treatment groups.

We may summarize the results reported to this point as follows:

1) When considering the overall main effects for scrambled as compared to ordered sequencing conditions, the results indicated that scrambled presentation of course materials produced more errors in the program, increased total instructional time, increased response time per individual question, and decreased the efficiency of instruction as indicated by the amount learned per unit of time. If one examined only the overall effects of sequencing on criterion performance without taking into account the differences in instructional time, and differences in student aptitude, nonsignificant differences were obtained. Results for the sequencing main effects indicated that the students made more errors during instruction in the scrambled sequence condition, but that they may have been able to compensate for the greater difficulty of the material by taking more time, by studying remedial frames, and by reorganizing relevant information as it was made available in the scrambled instructional sequence. Thus, by the end of instruction, the performance of students in the scrambled condition as measured by a criterion posttest was comparable to that of the students in the ordered condition.

2) The effects on student performance of scrambling a logically ordered instructional program seemed to depend in part on the ability of the learners. Three interactions (two of borderline significance) indicated that scrambling logically ordered course materials produced the greatest decrement in the performance of high-aptitude students even though these students spent the longest time in instruction. This effect was obtained for errors made in the program ($P < .10$), performance on the criterion measure ($P < .10$), and the efficiency measure ($P < .05$).

Discussion

Considered in the context of several previous investigations, the results of the pilot study suggest that the effects of scrambling a logically ordered instructional program depend on the nature of the learning task and on individual differences in the learners. Several previous investigations obtained no differences between scrambled and "ordered" versions of instructional programs, when the programs were relatively short, and when each frame in the program consisted of relatively discrete items of information. However, the results of the pilot study indicated that when a relatively long, complex program involving the learning of principles and problem solving was employed, item scrambling increased within program errors and instructional time but did not affect the final criterion level of achievement.

In the judgment of the writers, one of the most interesting tentative findings of the pilot study was that item scrambling impaired the performance of the high-aptitude students, but had little effect on the performance of the low-aptitude students. Although an aptitude by sequencing interaction had been predicted in the study, scrambled item sequencing was expected to have its most detrimental effect on the low-aptitude students. Just the reverse occurred. This finding seems most interesting in view of the frequently heard maxim that "highly able students can learn by almost any method."

What would account for the large decrements in the performance of the high-aptitude students in the scrambled sequence condition assuming that the effect is reproducible? Why did the low-aptitude students shown relatively little drop in performance in the scrambled sequence condition? One possible explanation for the results obtained is that one cannot impair performance by scrambling a program if performance is already quite poor. The large decrement in the performance of the high-aptitude students resulted in each case from the fact that their performance was quite high in the ordered condition and then dropped to the middle range in the scrambled sequence condition. The high-aptitude students had farther to drop. The low-aptitude students, on the other hand, started out in the middle range of the scale in the ordered sequence condition. This explanation might be plausible if there was evidence that the students were approaching the "floor" of the criterion test. However, the frequency distributions of performance on the criterion measure did not indicate that positively skewed distributions occurred which are typical when floor effects are encountered. If anything, the distributions tended to be slightly negatively skewed. Furthermore, the criterion measure was not a multiple-choice test; thus, a positive chance score on the test was virtually impossible. As was already indicated, most students achieved pretest scores of zero. In view of these characteristics of the measuring instrument, a mean score of 13 on the criterion test indicates that considerable learning occurred during instruction in the low-aptitude-ordered sequence group. In the writers' judgment, the failure of a performance decrement to occur in the low aptitude, scrambled sequence condition was not the result of floor limits in the measuring instrument or of the failure of the low-aptitude students to learn in the ordered sequence condition.

A more probable explanation of the finding that item scrambling produced larger decrements in the performance of the high-aptitude students than in the low-aptitude students was suggested by the results of the analysis of the student reaction data. The analyses of variance of the scales of the Student Reaction Inventory revealed significant effects for the three scales shown in Table 3. The interaction effect (again of borderline significance) for the Tense-Relaxed self-report rating scale was of special interest. The high-aptitude students in the scrambled sequence condition reported being more tense during instruction than each of the other three experimental groups. In addition, the high-aptitude students in the scrambled sequence condition reported the instruction to be "deeper" than the other three groups. This latter scale probably reflects the students' subjective perception of the difficulty of the program. Finally, the students in the scrambled sequence condition tended to rate the program as more inflexible than the students in the ordered condition. The students' self-reports on the Tense-Relaxed and Shallow-Deep scales suggest that the scrambled sequence program aroused the anxiety of the high-aptitude students. It is possible that the increased anxiety reported by the high-aptitude students in the scrambled sequence condition produced the decrement in their performance.

Table 3

Analyses of Variance of the Student Reaction Variables Tense-Relaxed, Shallow-Deep, Inflexible-Flexible for the High- and Low-Aptitude Students in the Scrambled and Ordered Sequence Conditions

		F-ratios		
Source	d.f.	Tense-Relaxed	Shallow-Deep	Inflexible-Flexible
Aptitude	1	.39	1.65	.04
Sequencing	1	2.26	.42	5.38**
Aptitude x Sequencing	1	3.92*	4.58**	.002
Error	43	(2.93)[a]	(1.31)[a]	(3.04)[a]

[a] Equals the mean square of the error term.
* P is less than .10
** P is less than .05

In view of the borderline P-values, the methodological problems encountered in the pilot study, and the fact that the direction of the differences were opposite to prediction, it is probably safest to withhold judgment on the nature of the aptitude by sequence interaction effect. Experiment I reported below was designed to provide a methodologically improved replication of the pilot experiment.

Scrambled versus Ordered Course Sequencing in Computer-Assisted Instruction: Experiment I

Experiment I was undertaken to replicate, extend, and improve upon the pilot study. There were several problems encountered in the pilot study:

1. Only approximately one-half of the Ss were assigned to the experimental treatments at random. Due to a number of programing problems which developed at the last minute in the scrambled sequence program, the random assignment of a large number of Ss had to be altered. Although there did not appear to be any selection factors in assigning Ss to treatments which could have affected the results of the study, we cannot be totally certain that these effects were negligible.

2. The modern mathematics CAI program employed in the pilot study contained remedial branching sequences. Thus, not all the Ss received identical instructional treatments. Those Ss with high error rates were more likely to be sequenced through remedial material than Ss with low error rates. Although branching programs are most appropriate for computer-assisted instruction, they present some obvious problems when one wants to make comparisons of experimental treatments which are unconfounded with the effects of remedial branching. The results of the pilot study suggested that the Ss in the scrambled sequence condition might have benefited from the greater number of remedial frames which they received. Although the scrambled sequence Ss made significantly more errors in the program, their performance at the end of instruction was not significantly different from that of the ordered sequence Ss as measured by an achievement posttest.

3. The CAI instructional program employed in the pilot study was in its preliminary stages of development. This original program had not been systematically evaluated or revised, and although student criterion test performance on the ordered version of the program was adequate for a first draft program (see Figure 1, page 11) it was decided that further experiments should be conducted on an improved version of the modern mathematics instructional program.

The major objectives of Experiment I were essentially the same as those of the pilot investigation. The major changes were methodological in an attempt to avoid the difficulties encountered in the preliminary experiment. The specific objectives were as follows:

a) To determine the effects on student learning of scrambling a logically ordered instructional program when a relatively long and difficult program is used, and when the subject matter involves a concept hierarchy. A section of a modern mathematics program on converting base ten numbers to non-decimal base numbers appeared to meet these criteria and was chosen as the learning task for the investigation.

b) A second major purpose of the study was to examine a number of interactions between several student individual difference measures, and the experimental treatment (scrambled versus ordered sequence). An interaction of marginal statistical significance was obtained in the pilot study indicating that high aptitude *Ss* as measured by the *Scholastic Aptitude Test* were more detrimentally affected by the scrambled sequence of instruction than were low aptitude *Ss*. This finding was inconsistent with the prediction that a scrambled instructional sequence would be most detrimental to the learning of low aptitude *Ss*. Experiment I served as a replication of this rather surprising, but tentative finding.

Methods and Procedures

Revision of the CAI instructional program

The instructional program was totally revised by Harold Sands of the Penn State CAI Laboratory. The original criterion achievement test (see Appendix B) defined the objectives of instruction. The instructional program was revised in the hopes of producing a more effective program as measured by criterion test performance. The following subsets of topics were included in the revised program: 1) A review of working with exponents and raising a number to its power, 2) Expand a number in any number system and understand the concept of place value, 3) Convert a number in any base to its base ten equivalent using an appropriate conversion algorithm, and 4) Convert a number from base ten to any other base using an appropriate conversion algorithm. Program samples of the original and revised programs are shown in Appendix A. Preliminary evaluation of the revised program indicated that it would be suitable for the experiment, and that it produced some improvement in student learning as measured by the criterion test over that obtained with the original program.

One major change in the instructional program consisted of the elimination of the remedial branching segments and the development of the course as a linear program. As indicated above this change was necessitated by the requirement that the experimental treatments be identical in all respects except the sequence of instruction. Since the branching program adapted automatically to the performance of individual Ss, it was extremely difficult to insure identical instruction in the two experimental groups.

The revised program consisted of 74 instructional frames and took, on the average, two hours for the students to complete. That the program was reasonably satisfactory in accomplishing the instructional objectives can be seen by examining the distribution of performance of 37 college Ss on the criterion test shown in Figure 4. All but five of these Ss obtained scores of zero on a pretest measure of ability to make the number system conversions. The distribution of criterion test performance is negatively skewed about a mean of 17.03 out of a possible 22 total points on the criterion test (one item was deleted from the original 23-item test used in the pilot experiment). Thus, although a small number of Ss made only modest progress in the course, the majority of the Ss appear to have mastered the relatively difficult task of converting numbers into nondecimal bases.

The number systems course was programed for presentation to the Ss via CAI using the Coursewriter language (Maher 1964) developed by I.B.M. The decision logic of the Coursewriter language enabled us to provide either an ordered or scrambled sequence of frames from the same basic materials. This was accomplished by the use of conditional branch statements following each frame. When a student signed on to the program he typed either "S" for scrambled or "U" for unscrambled depending on the experimental treatment to which he had been randomly assigned. At this time the computer stored a certain value in a counter which identified the subject as in the "s" group or the "u" group. Following each problem in the program, the computer checked this value. If the computer identified the subject as an "s" subject, it presented the next frame in a pre-established random sequence (determined from a table of random numbers). If the subject was identified as a "u" subject the computer presented the next frame in the ordered sequence. A flow-chart for one frame illustrating this simple programing expedient is shown in Figure 5.[1] This strategy avoided the necessity of having to develop two completely separate versions of the program.

---

[1]The writers wish to acknowledge the contribution of Dr. David Gilman of the Penn State CAI Laboratory for suggesting this programing strategy.

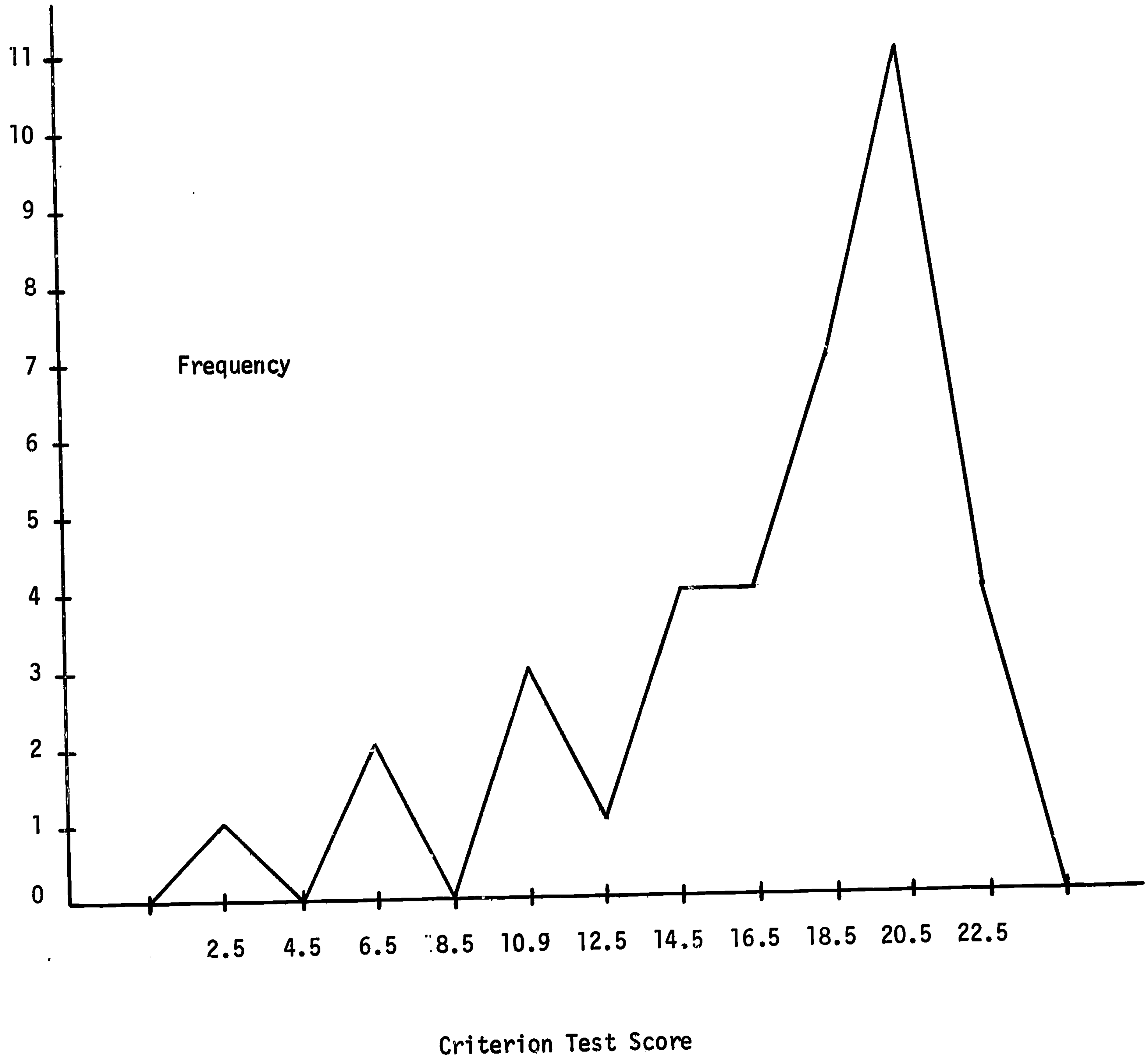


Fig. 4. Frequency distribution of the criterion test performance of 37 Ss in the ordered sequence.

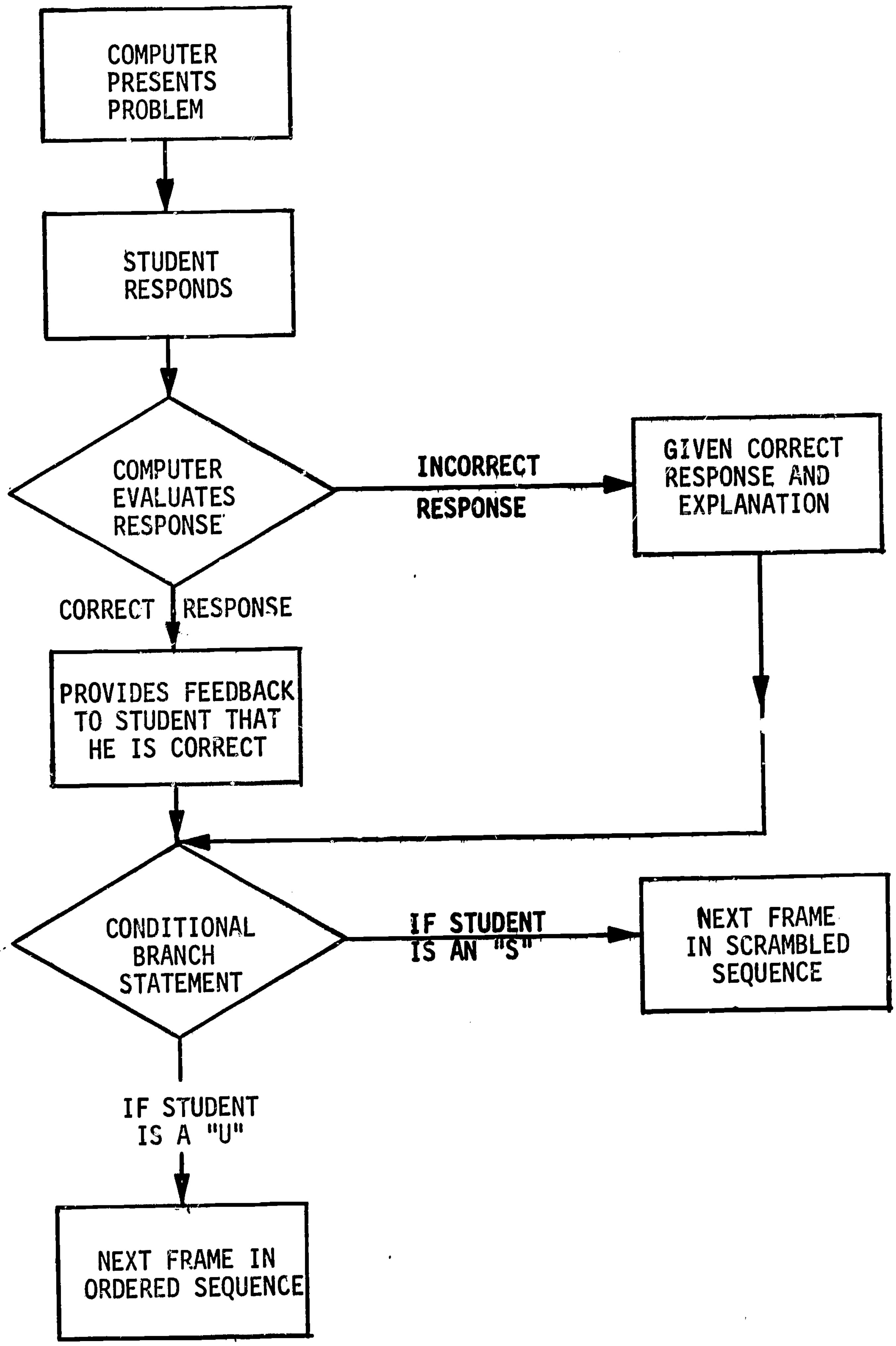


Fig. 5. Flow-chart of an instructional frame

Description of the CAI System

The central computer used in Experiment I was an IBM 1410 system located in the computation center on the Penn State campus. The reader will recall that in the pilot study, the experimental course materials were teleprocessed to student terminals on the Penn State campus from an IBM computer located in Yorktown Heights, New York. The CAI system employed in Experiment I was located entirely on the Penn State campus. Four IBM 1050 typewriter terminals, located in the Penn State CAI Laboratory were employed in the investigation. The course was typed out to the Ss at the terminal, and students entered their responses by typing them in at the terminal.

Subjects and Procedures

Eighty undergraduate students at the Pennsylvania State University served as the Ss in the experiment. Ss were enlisted from a large section of introductory educational psychology and a large section of a course in instructional media. The subject population consisted predominantly of education majors. All Ss participated in the experiment during the fall term, 1966. Participation in the experiment was strictly voluntary, and each S who completed all phases of the experiment received an honorarium of $4.00. The only requirement for participation in the study was that the student had not had previous course work in non-decimal number systems, and that his performance on the pretest indicated little or no prior knowledge of the subject matter. Ss were assigned at random to either the scrambled or ordered sequence condition. At the completion of the experiment, the number of Ss for whom complete data were available on the major dependent variables of the study were 41 in the scrambled sequence condition and 34 in the ordered sequence condition.

Ss reported to the CAI Laboratory individually and were met by a proctor. Four proctors were used in the study, one for each of the four student terminals. The following standard directions were read individually to each S by a proctor:

> You see in front of you a two-way typewriter commuication device used for computer-assisted instruction. A computer presents problems to students over the typewriter, and the student enters his answer by typing it on the machine. In a moment you will be shown how to operate the typewriter. We are engaged in an experiment to test the feasibility of teaching by computer. In order to get a fair evaluation of the CAI (computer-assisted instruction) method, it is very important that you do the best job you can on the course which you will take. We also must

know how much knowledge you already have in the subject matter to be taught. Have you ever had Math 200? Have you had any previous courses in high school or college dealing with number systems with bases other than ten? (If student says "yes" explain that he has already had the material to be taught, that we can use him in a later experiment, and excuse him from the experiment). First, you will be given a pretest to measure your achievement in the subject. Following instruction you will be given two more tests. One test will be taken immediately following instruction today, and the other will be taken tomorrow at a scheduled room. Your scheduled room and time to take the test tomorrow is __________. (Make sure student writes this down). Tomorrow you will also be given a short questionnaire to get your reactions to computer-assisted instruction. Remember in order to obtain the fee for participation in this experiment, you must complete the entire experiment.

Before we begin, I want to get some general information from you. (Fill out file card on student).

I want to see how you can do on this pretest, just to check on your knowledge in this subject. I know that you probably have not had instruction on this topic, and that you may not be able to do many of the problems. Don't worry about that. Just do the best you can. If you cannot do a problem just leave it blank. I will return in a few minutes to see how you are doing.

Before we begin the actual instruction, there are a few more things you may want to know. Although the time to complete the course varies depending upon the particular student, students on the average have taken about two hours to finish the course. After the first hour there will be a five minute break. The Proctor will notify you when it is time to take the break, and when it is time to start working again. There are two handouts you will need while working through the program. At several points in the course the typewriter will refer you to certain exhibits. These are the exhibits you will need. Be sure to refer to the exhibits at the appropriate time. In addition, to the exhibits, there will be times when the computer will ask you to work out a problem using paper and pencil, and then compare your solutions with the solution typed out by the computer. When you are asked to solve a problem

> using paper and pencil show all your work and the answer in the appropriate space on this work sheet. As you work through the problems you are not to look back at previous problems in the course.
>
> If you suspect any difficulty with the machine, call on the proctor for assistance. However, the proctor has been instructed to give you no help with the problems or the content of the course. Do you have any questions before we begin?

After the reading of the directions the S was given an opportunity to solve the problems on the pretest. Most Ss had no idea how to solve the pretest problems. Ninety per cent of the Ss obtained pretest scores of zero, and only two Ss obtained pretest scores greater than one.

Following the pretest, Ss were given a brief explanation of how to operate the typewriter terminal, and the proctor assisted in signing the student on to the appropriate experimental version of the program. The proctor remained with the S through the first few frames of the program to insure that he had learned how to enter his responses at the terminal. The S was then left to work through the program at his own rate. A five-minute rest period was included approximately half-way through the instruction. At the completion of instruction the Ss were administered an alternate form of the pretest to measure their ability to perform the number systems conversions. The following day, the Ss returned to the CAI Laboratory to take a transfer test and the Student Reaction Inventory measure of attitude towards CAI. In a few cases a weekend or several days intervened between instruction and the administration of the transfer and attitude measures, however, in no instance was the interval between instruction and the transfer measure greater than 4 days.

In summary, the independent variables of the investigation consisted of one manipulatable treatment effect (scrambled versus ordered sequence of instruction), four measured individual difference variables (verbal SAT, quantitative SAT, total SAT, and grade point average), and sex. The dependent variables consisted of the number of errors made in the program, time taken to complete the program, scores on the achievement posttest, transfer test scores, and attitude towards CAI as measured by the Student Reaction Inventory. The latter three measures are reproduced in Appendix B. As reported above, the test-retest reliability of the achievement posttest was found to be .93. Since the transfer test was found to be exceedingly difficult for the Ss, its reliability could not be adequately assessed in the present study. Brown (1966) obtained a Hoyt reliability of .88 for the Student Reaction Inventory.

Results

The results were analyzed by means of single-factor analyses of variance and within groups correlations. The analyses of variance provided information on the overall effects of the scrambled versus ordered sequence on student learning and transfer. The within groups correlational analysis provided information on the interactions between the student individual difference measures and the experimental treatment.

The main effect of program sequence on student learning

Means, standard deviations, and the analyses of variance summaries are shown in Table 4 for the instructional time and error measures. These results are consistent with the results obtained in the pilot experiment. Scrambling the sequence of instruction increased instructional time ($P < .10$) and increased the number of errors made in the program ($P < .001$). However, when one examines the analyses of the achievement posttest and transfer measures shown in Table 5, one finds nonsignificant differences between the scrambled and ordered sequence conditions. This result is again consistent with the results of Experiment I (see Table 1, page 11). In the pilot experiment, the Ss in the scrambled sequence condition made more errors, and took longer to complete the program than the Ss in the ordered sequence condition, but by the end of instruction the scrambled sequence Ss were about on a par with the ordered sequence Ss as measured by the achievement posttest. In Experiment I, there were again differences between the treatment groups on errors and instructional time, but no significant differences on the achievement posttest administered at the termination of instruction. These data support the conclusion that although the scrambled sequence is detrimental in terms of errors produced and instructional time, Ss are able to overcome the difficulties produced by the scrambled sequence and eventually arrive at a level of criterion performance comparable to the ordered sequence group.

Further light is shed on this finding by examining the errors made per block of ten frames by the two treatment groups as shown in Figure 6. The ordered sequence group shows a gradual increase in the number of errors over the course of the program beginning at a mean of 1.2 errors persubject in the first ten frames, and rising to a mean of 4.1 in the last twelve frames. This increase in errors for the ordered group simply reflects the gradually increasing difficulty of the problems in the program. While early frames involved relatively simple problems such as raising a number to its power, and expanding a number, the program builds in difficulty to criterion level problems involving the conversion of a number in one base, to its equivalent in another base. Problems

Table 4

Summary Statistics and Analysis of Variance
for the Time and Error Scores

Dependent Variable	Scrambled Sequence (n=43) $\bar{X}$	s.d.	Ordered Sequence (n=37) $\bar{X}$	s.d.	Analysis of Variance: Mean Squares Treatment (1 d.f.)	Within (78 d.f.)	F	P
Total Instructional Time (minutes)	130.70	26.80	121.00	20.40	1877.80	579.70	3.24	<.10
Errors	16.30	5.68	11.60	5.40	422.13	30.83	13.69	<.001

Table 5

Summary Statistics and Analysis of Variance
for the Achievement Posttest, Transfer Test,
and Attitude Measures

Dependent Variable	Scrambled Sequence (n=41) $\bar{X}$	s.d.	Ordered Sequence (n=34) $\bar{X}$	s.d.	Analysis of Variance: Mean Squares Treatment (1 d.f.)	Within (73 d.f.)	F	P
Achievement Posttest	18.50	9.40	17.10	5.00	36.40	59.80	< 1.0	
Transfer	5.29	4.40	4.76	4.56	5.18	20.00	< 1.0	
Attitude Towards CAI	77.10	9.89	74.60	12.40	113.80	122.90	< 1.0	

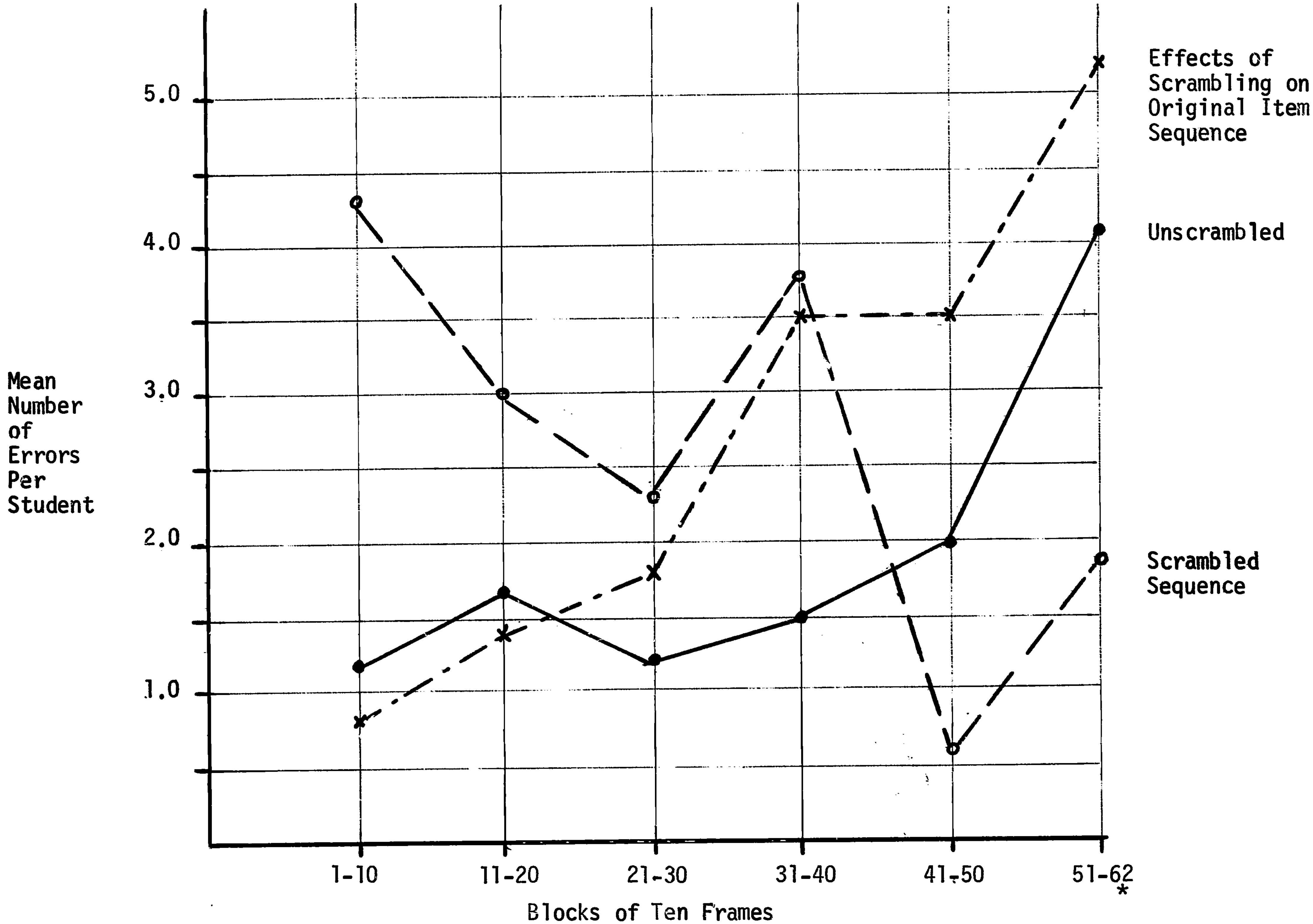


Fig. 6. Mean number of errors per block of ten instructional frames for the two treatment groups. (scrambled n = 43; ordered n = 37)

*Twelve frames in this group.

in the later stages of the program require that the student integrate and apply what he has learned in the early stages of the program. The error curve labeled "effects of scrambling on original item sequence," shows the effects of scrambling on the errors produced to frames as they were originally sequenced in the ordered program. This curve starts at the same point as the performance curve for the ordered sequence group, but eventually diverges to a considerably higher error rate in the middle and later frames. The comparison of these two performance curves supports the view that the original "ordered" program did in fact contain sequentially dependent frames. Thus, when frames 31-40 are taken out of their original ordered sequence and scattered at random throughout the program the mean number of errors produced to these frames increases from 1.5 to 3.5. When frames 51-62 were scrambled, the mean number of errors increased from 4.1 to 5.2. The ability to solve problems in the middle and later stages of the program clearly depends to some extent on knowledge acquired in the earlier stages of the program. When the original sequence was altered, the number of errors increased. Perhaps the most interesting performance curve shown in Figure 6 is the one simply labeled "scrambled sequence." This curve shows the errors by blocks of ten frames as they were actually encountered in the scrambled program. The error rate starts high as a mean of 4.3 in the early frames as the student encounters problems for which he still does not have the prerequisite knowledge. As more and more of necessary prerequisite knowledge is pieced together in the scrambled program, the errors gradually decrease, until in the last twenty frames, the Ss in the scrambled sequence are making fewer errors than those in the ordered condition. Although scrambling an instructional program in which there is an inherently ordered sequential dependency among the frames does increase the number of errors in the early stages of instruction, Ss are eventually able to organize the necessary information so that by the end of instruction they are performing at a criterion level commensurate with that of students taught by the ordered sequence.

Clarification of what may be operating in this study may be aided by a simple hypothetical example. Suppose that a certain learning task involves the mastery of four concepts A, B, C, and D, and the application of the four concepts to the solution of a problem called E. The four concepts form a naturally ordered sequence such that the learning of B depends upon an understanding of A, the learning of C depends upon an understanding of A and B, and so on. The learning task is now taught to one group of students in the naturally ordered sequence A, B, C, and D, and to another group in the sequence D, C, B, and A. The latter group of course will make more errors during instruction. Questions pertaining to concept D cannot be answered correctly because the prerequisite concepts C, B, and A have not yet been mastered. It is highly probable that immediately after instruction on concept D, that concept D is still not in a learned state for most of the Ss. As the Ss receive exposure to concepts C, and B more of the relevant information becomes available but complete understanding may still be lacking for want of concept A. Finally, instruction is completed and the Ss have been exposed to all of the relevant concepts. Now that concept A is available, concept B may be understood, and once concept B is available, concept C may be understood, and when concepts A, B, and C are understood then concept D becomes clear. This interpretation suggests that adult Ss are able to relate the relevant information as it is made available by the program to the previously poorly understood concepts which were presented out of sequence. The major requirement of such an interpretation would be that the Ss were capable of recalling relevant aspects of previous problems at some later point in the program when the information needed to solve them became available. An additional requirement would seem to be that the Ss have the ability to recognize the relevance of material provided later in the program to problems which were presented earlier. This interpretation leans heavily on the cognitive processes of the Ss which allow them to integrate and organize information regardless of the sequence of presentation.

Analysis of the measure of attitude towards CAI produced nonsignificant differences between the scrambled and ordered sequence groups. In addition, there did not appear to be any sex differences in performance on the dependent variables and no sex by treatment interactions.

Treatment by aptitude interaction effects

Two methods were used to examine the interactions between aptitude as measured by the SAT and instructional sequence. Following the method discussed by Cronbach (1957), correlations were computed between the aptitude measures and the various dependent variables within each of the two experimental treatment conditions. The interaction between a measured individual difference variable such as aptitude and a manipulatable experimental treatment can be assessed by comparing the correlations of the measured ID variable with the dependent variable within each level of the manipulatable experimental treatment. If these within groups correlations are not significantly different, the interaction between the measured independent variable and the manipulatable treatment condition can be considered nonsignificant. If the within groups correlations are significantly different indicating the presence of an aptitude by treatment interaction effect, one can proceed to plot the regression lines of the dependent variable on the measured independent variable within each of the experimental treatment conditions to examine the nature of the interaction effect. This method was suggested by Cronbach (1957) as a method for bridging the gap between the two traditional disciplines of psychology: the experimentalists and the individual difference correlational psychologists. Cronbach was advocating an approach to experimentation which examined interactions between measured individual difference variables and manipulatable experimental treatments, an approach consistent with the sympathies of the present authors.

The approach to examining interactions between experimental treatments and measured individual difference variables described by Cronbach, may be contrasted with a somewhat more familiar approach to examining such interactions. In the typical study which involves both a measured independent variable such as test anxiety, aptitude, etc., and an experimental treatment, the Ss may be divided at the median into high and low anxiety or high and low aptitude groups. The data are then analyzed in the form of a traditional factorial analysis of variance design which yields the somewhat more familiar interaction between the individual difference variable and the experimental treatment. Although the latter procedure is the more common of the two methods, the assessment of interaction effects by means of within groups correlations may be the more preferred procedure since it avoids arbitrary splits of a continuous score distribution into high and low groups, and treats what is essentially a correlational problem as a correlational problem examining the interaction over the entire range of the measured independent variable. The factorial analysis of variance approach may be the preferred method when curvilinear relationships exist between the measured independent variable and the dependent variable. In the present experiment, the writers decided to examine the interactions between aptitude as measured by the SAT and course sequencing by means of comparisons of the within groups correlations. In addition, to aid in the interpretation of the results, several of the more common factorial analyses of variance were computed by dichotomizing the aptitude distributions into high and low groups.

The correlations between the SAT scores, college grade point average (GPA), and the dependent variables computed separately for each of the two treatment groups are shown in Table 6. A good many of the correlations with the SAT scores differ significantly from zero and the directions and general magnitude of the relationships are consistent with expectations. The key comparisons for examining the aptitude by sequencing interactions involve the comparison of the correlations of an independent variable (SAT score) with a dependent variable in the two treatment conditions. For example, an interaction would be evident if the correlation between the SAT verbal and the achievement posttest in the scrambled condition was significantly different from the correlation between these same two variables in the ordered sequence condition. The correlation between the SAT

Table 6

Correlations Among the Dependent and Independent Variables Within Each of the Two Treatment Groups*

Independent Variables	Dependent Variables: Scrambled Sequence					Ordered Sequence				
	Achievement Posttest	Transfer	Errors	Time	Attitude	Achievement Posttest	Transfer	Errors	Time	Attitude
SAT (Verbal)	.57	.35	-.72	-.48	.27	.38	.52	-.46	-.39	.08
SAT (Quantitative)	.74	.44	-.48	-.24	.13	.58	.36	-.60	-.49	.08
SAT (Total)	.80	.48	-.74	-.44	.24	.57	.51	-.63	-.51	.10
GPA	-.08	-.03	.02	-.14	.09	.08	-.07	-.17	-.42	-.09

*N's for correlations involving SAT scores: Scrambled sequence n = 21, Ordered sequence n = 22, $r_{.05} = .43$

N's for correlations with GPA: Scrambled sequence n = 41 ($r_{.05} = .31$), Ordered sequence n = 37 ($r_{.05} = .32$)

verbal and the achievement posttest in the scrambled sequence condition was .57 while the correlation between these same two variables in the ordered condition was .38. Each correlation was transformed via Fisher's Z-transformation and the difference between the two was tested for statistical significance following the procedure outlined by McNemar (1962, page 139). The difference between the correlations of .57 and .38 was not statistically significant. Each correlation in the scrambled sequence condition was compared with the correlation between the same two variables in the ordered condition. None of the pairs of correlations was significantly different, indicating that the aptitude by sequencing interactions were nonsignificant as assessed by the correlational analysis.

In addition to the statistical nonsignificance of the differences between the correlations reported in Table 6, it is important to note that the direction of the differences is not consistent with the interactions obtained in the pilot study. The reader will recall that in the pilot study the high aptitude Ss were more detrimentally affected by the scrambled sequence than the low aptitude Ss. In order for this finding to have been replicated in Experiment I, the correlations between the SAT scores and the dependent variables would have to have been larger in the ordered sequence condition than in the scrambled sequence condition. On the contrary, the correlations indicated that where the largest differences occurred, the correlation was largest in the scrambled sequence condition rather than in the ordered condition. Thus, the correlations between the SAT verbal, quantitative, and total scores, and the achievement posttest were .57, .74, and .80 respectively in the scrambled sequence condition, but the correlations between these same measures in the ordered condition were .38, .58, and .57 respectively. The correlation between SAT verbal and errors was -.72 in the scrambled sequence condition, and -.46 in the ordered sequence condition. These results, although nonsignificant statistically, are consistent with the results obtained by Stolurow (1964), and are consistent in direction with the original hypothesis of the present series of investigations. The writers' original expec- was that scrambling the sequence of instruction would tend to increase the covariance between measures of aptitude and performance during instruction.

The aptitude by sequencing interactions were also examined by means of the more traditional factorial analysis of variance approach, Ss in each of the two sequence groups for whom Verbal SAT scores were available were dichotomized at the median into high and low aptitude groups. Table 7 shows the results of a 2 x 2 factorial analysis of variance with errors as the dependent variable. The effect of major interest here is the aptitude by sequence interaction which was statistically significant at less

Table 7

Analysis of Variance of Errors for High and Low Verbal SAT Groups in the Scrambled and Ordered Sequence Conditions

Source of Variation	d.f.	M.S.	F.	P
Sequence	1	581.2	22.8	$<.001$
Aptitude	1	386.9	15.2	$<.001$
Sequence x Aptitude	1	108.3	4.26	$<.05$
Error	39	25.4		

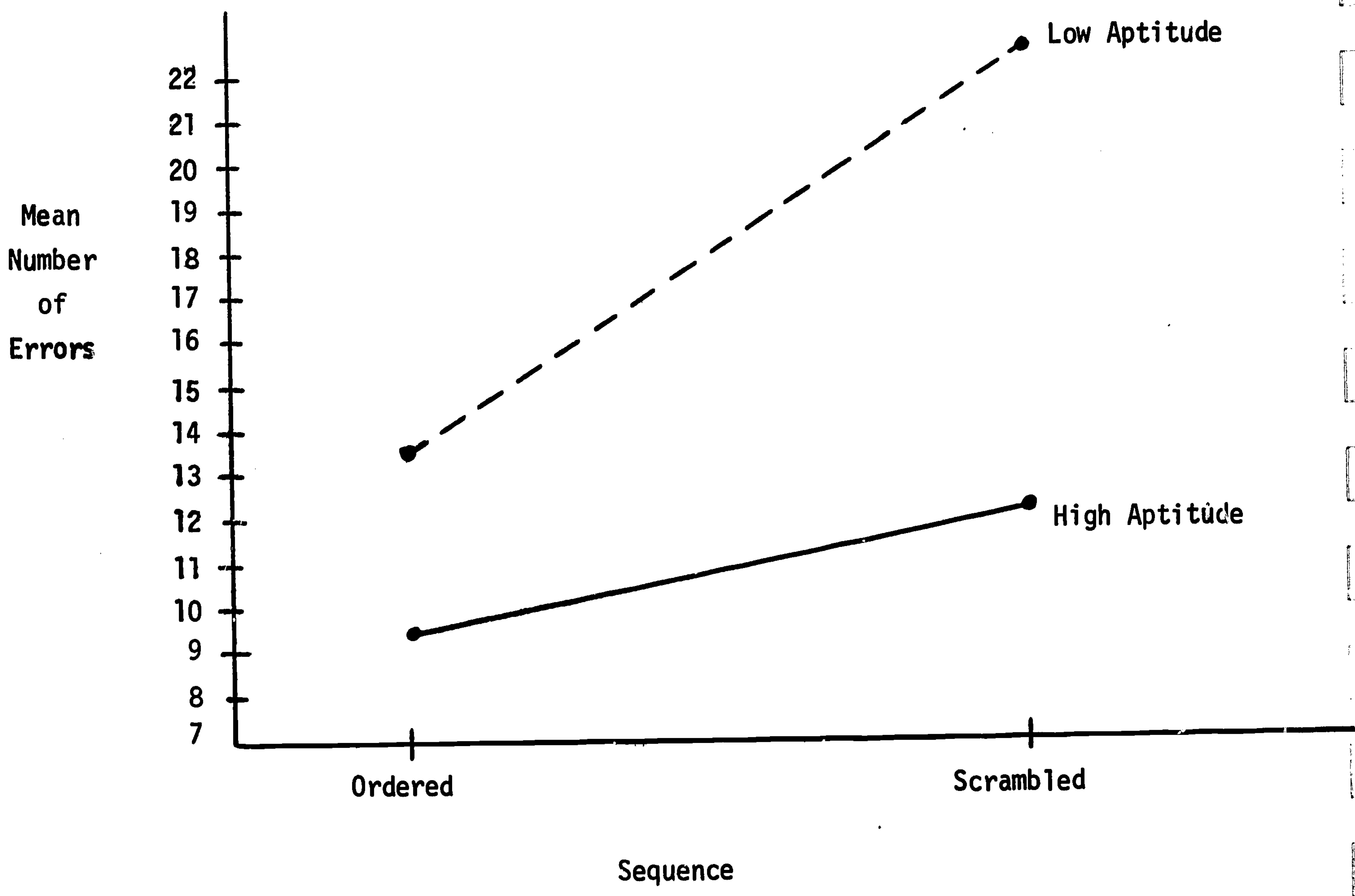


Fig. 7. Mean number of errors of high and low aptitude Ss taught by scrambled and ordered programs (N's equaled: HA-ordered = 11, HA-scrambled = 10, LA-ordered = 11, LA-scrambled = 11)

than the .05 level. This interaction has been plotted in Figure 7 which can be compared with the plot of the same interaction in the pilot study (see Figure 2 on page 12). The significant interaction shown in Figure 7 indicates that the scrambled sequence had the most detrimental effect on the performance of the low aptitude Ss. This finding is consistent with the original hypothesis of the study, but inconsistent with the findings of the pilot study where the scrambled instructional sequence appeared to have its most detrimental effect on the performance of high aptitude Ss. The discrepancy between the two studies does not appear to be explainable in terms of differences in the abolute levels of aptitude of the groups. In the pilot study the mean of the high verbal SAT group was 612 and that for the low group was 435. In Experiment I, the mean of the high group was 612 and the mean of the low group was 477.

The factorial analysis of variance with total instructional time as the dependent variable is shown in Table 8. The aptitude by sequence interaction was statistically significant at less than the .05 level. The plot of this interaction shown in Figure 8 indicates that the scrambled sequence condition increased instructional time for the low aptitude Ss, but produced a slight decrease in instructional time for the high aptitude Ss. This

Table 8

Analysis of Variance of Total Instructional Time for High and Low Verbal SAT Groups in the Scrambled and Ordered Sequence Conditions

Source of Variance	d.f.	M.S.	F	P
Sequence	1	597.17	2.59	
Aptitude	1	3566.13	15.45	<.001
Sequence X Aptitude	1	1178.11	5.11	<.05
Error	38	230.74		

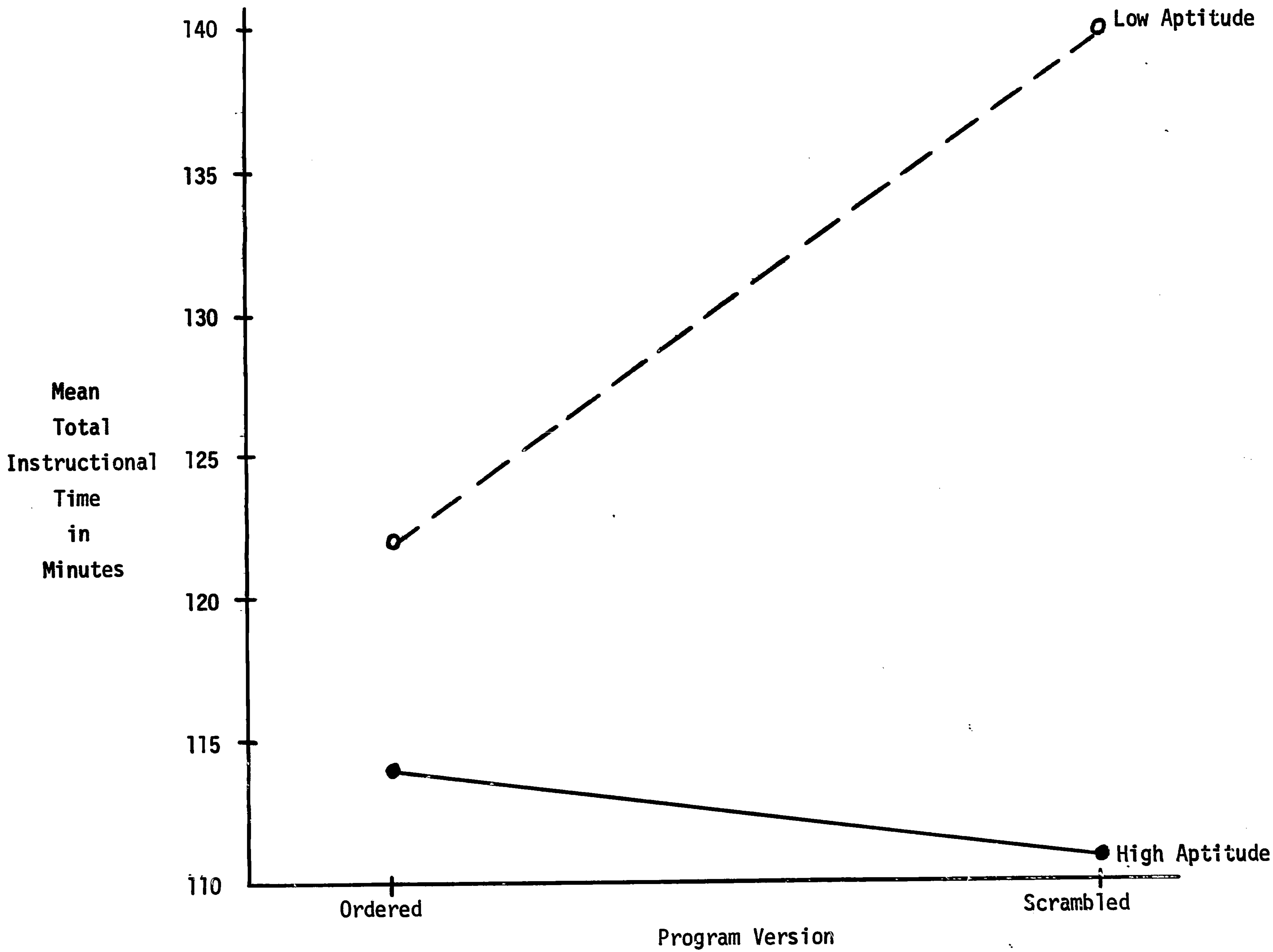


Fig. 8. Mean instructional time of high and low aptitude Ss taught by scrambled and ordered programs (N's equaled: HA-ordered=11, HA-Scrambled=10, LA-Ordered=11, LA-Scrambled=11)

finding is again consistent with the original hypothesis of the study, that the performance of low aptitude Ss would be more detrimentally affected by a scrambled sequence of instruction than the performance of high aptitude Ss, however, it is inconsistent with the results of the pilot study.

Finally, the factorial analysis of variance with the achievement posttest as the dependent variable is shown in Table 9. Here the aptitude by sequence interaction is statistically non-significant ($F=2.70$, $P > .10$), however, again the direction of the differences is inconsistent with the results of the pilot study, but consistent with the other interactions obtained in Experiment I. The plot of the aptitude by sequence interaction shown in Figure 9 indicates that the scrambled sequence produced a slight increase in the posttest performance of the high aptitude Ss, but a slight decrease in the performance of the low aptitude Ss. The reader may wish to contrast the interaction plotted in Figure 9 with the interaction obtained for the same dependent variable in the pilot study as plotted in Figure 1 on page 11. The direction of the interaction obtained in Experiment I is again consistent with the original hypothesis that the sequence of instruction interacts with Ss' aptitude such that a scrambled sequence is more detrimental to the performance of low aptitude Ss than high aptitude Ss.

Table 9

Analysis of Variance of Posttest Scores for
High and Low Verbal SAT Groups
in the Scrambled and Ordered
Sequence Conditions

Source of Variance	d.f.	M.S.	F	P
Sequence	1	1.65	< 1.0	
Aptitude	1	219.06	9.90	$< .001$
Sequence x Aptitude	1	59.73	2.70	
Error	39	22.12		

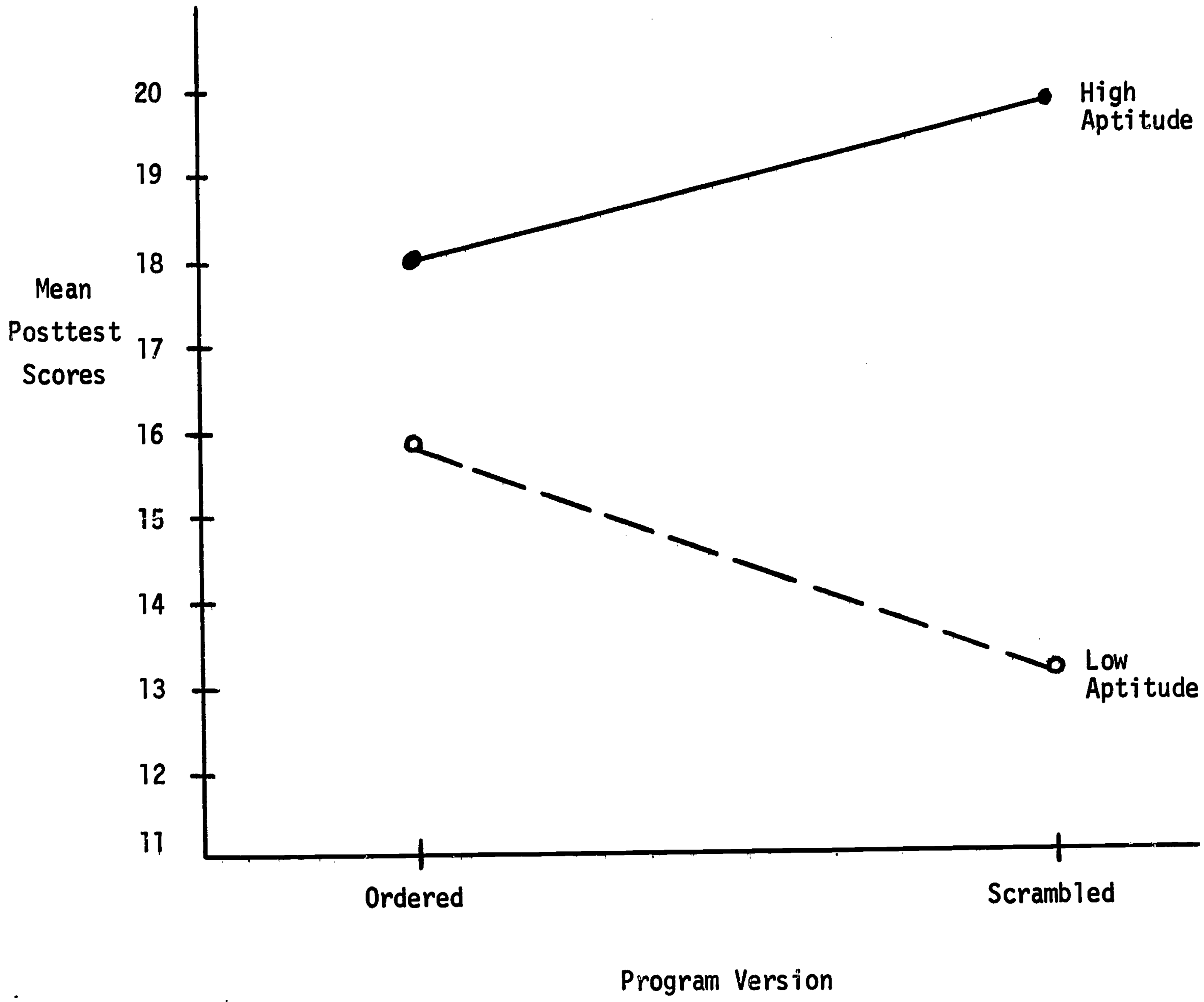


Fig. 9. Mean Posttest Performance of High and Low aptitude Ss taught by Scrambled and Ordered Programed (N's equaled: HA-ordered = 11, HA-scrambled = 10, LA-ordered = 11, LA-scrambled = 11)

In summary, the analysis of the aptitude by sequencing interaction effects by the two methods of comparing within groups correlations, and the factorial analysis of variance produced results which although consistent in direction, were not consistent with regard to statistical significance. In general, the correlational analysis produced nonsignificant interactions, while the factorial analysis of variance produced several interactions which reached acceptable levels of statistical significance. These discrepancies between two methods of analyzing the same data raise some interesting questions concerning the appropriateness of these two procedures for assessing interactions between individual difference measures and experimental treatments. For example, it would be particularly interesting to determine the relative power of these two statistical methods for detecting "true" interaction effects. In view of the importance of the problem of assessing interactions between individual difference variables and experimental treatments the two senior authors are presently conducting a methodological investigation of this problem.

In view of the inconsistency of the results of the pilot study and Experiment I concerning the nature of the interaction between student aptitude and instructional sequence, one must exercise caution in reaching any conclusions concerning this phenomenon from these two experiments. The first preliminary study produced an unexpected and rather unusual interaction at a borderline level of statistical significance suggesting that the scrambled sequence of instruction was most detrimental to the performance of high aptitude Ss. The second study produced two significant aptitude by sequencing interactions for errors and total instructional time with the direction of the effect opposite to that of the pilot experiment but consistent with the original predictions of the study. In Experiment I the scrambled sequence of instruction was most detrimental to the performance of the low aptitude Ss on errors and instructional time, and in the same direction, but not significantly so for achievement posttest. These discrepancies between the two studies may have resulted from the partially nonrandom assignment of Ss in the pilot study (resulting in the operation of an unknown selection factor), or the change in the instructional program resulting from the revision of the program prior to Experiment I.

Scrambled Versus Ordered Course Sequencing in Computer-Assisted Instruction: Experiment II

Previous investigations produced inconsistent findings concerning the importance of ordered sequencing in programed instruction. Evans (1960) and Stolurow (1964) reported superior learning in carefully organized instructional programs, while Roe, Case, and Roe (1962) and Levin and Baker (1963) reported nonsignificant differences between scrambled and ordered sequence conditions. One experiment (Hamilton, 1964) reported differences in student learning favoring a scrambled sequence condition. Since these previous investigations employed different subject matters, one possible explanation for the inconsistent findings may lie in the nature of the different subject matters employed. Some subject matters may involve a natural hierarchy of concepts and a sequential dependency among successive stages of the course. Indeed, the number systems program which was used in Experiments I was specifically selected because it appeared to have such hierarchical characteristics, and the performance data reported in Figure 6 bear out this conclusion. The frequency of errors increased when the original sequence of frames was altered. The expectations of the investigators was that scrambling such an inherently ordered subject matter would have a highly detrimental effect on student learning. Somewhat contrary to these expectations, the results of Experiments I and the pilot study indicated that although the scrambled sequence led to a greater number of errors made during instruction, and to an increase in instructional time, by the end of instruction the Ss in the scrambled condition were apparently able to master the material since their posttest problem solving was comparable to that of the ordered sequence Ss.

Not all subject matters contain the conceptual structure inherent in such fields as modern mathematics. Performance in subject matters consisting of a set of relatively discrete facts, such as the learning of vocabulary, knowledge of terms, and anatomy, would not be expected to be detrimentally affected by varying the sequence of presentation. Experiment II was undertaken to provide information on the effects of a scrambled sequence on the learning of such a subject matter. In this phase of the study we were primarily concerned with the influence of subject matter characteristics on the effect of course sequencing. By extending Experiments I to another subject matter it was possible to test the generalizability of the effects obtained with the modern mathematics program. In addition, by comparing the results of the three experiments we hoped to gain some insight into the critical characteristics of subject matters which might interact with the sequence in which the material is taught.

Method

Description of the CAI instructional program

Part of a CAI program written by Professor Bruce Siegenthaler and Mr. Jeffrey Katzer of the Penn State CAI Laboratory was adapted for the present experiment. A more complete description of the course may be found in an earlier report by Mitzel (1966). The first section of the course entitled Speech Pathology and Audiology was adapted for the study. The course segment consisted of instruction on the anatomy of the ear. The student's task was to learn the anatomical names of the parts of the ear. (A sample segment of this program is shown in Appendix B.) Unlike the number systems program employed in the first two studies, the anatomy of the ear program did not appear to involve any inherently logical sequence. On an a priori basis, it did not appear to make much difference whether the student learned one anatomical part before or after another anatomical part. The task appeared to consist of a set of relatively discrete associations, and closely resembled a paired-associates learning task in which the stimulus item consisted of a part of the ear, and the response item consisted of the name of the part. The anatomical names were extremely rare in everyday usage thus, we could be relatively certain that unless the Ss had completed previous course work in the area, they could not make the appropriate responses. (Preliminary research with this program supported this assumption.)

The instruction was presented to the Ss by means of an IBM 1410 CAI system which was identical to the system used in experiment I. The student-subject matter interface consisted of an IBM 1050 typewriter terminal through which questions were displayed to the student and through which the student could enter his responses. In addition to the portion of the course which was administered via the typewriter, Ss had available at the terminal a plastic model of the ear and several handout sheets containing line drawings of the ear. At several points in the program, the students were asked to identify a certain anatomical part on the model of the ear, and on the handout sheet, and to write the name of the appropriate part on the handout sheet.

Sixty-four undergraduate students in an introductory educational psychology class and in a course on instructional media served as the experimental Ss. The Ss volunteered for the experiment and received an honorarium of either $2.00, $3.00, or $4.00 depending upon the average length of time required to complete the experiment by the treatment group to which they had been assigned. The Ss were assigned at random to one of six treatment conditions in a 2 x 3 factorial analysis of variance design. One main effect consisted of the scrambled versus ordered sequence of instruction. The second main effect consisted of the number of times that the student was cycled through the program. The basic anatomy of the ear program contained 22 instructional frames. Preliminary tryouts of the program indicated that most students could not recall very many of the anatomical names after only one trial through the program. For this reason it was necessary to vary the number of trials through the program so that the effects of the scrambled sequence could be observed over a number of practice levels. The main effect for trials consisted of three levels: one, two, and three trials through the program. This design also permitted the examination of the interaction between the sequencing variable and the amount of practice. It seemed probable that the effects of the scrambled sequence might be offset by increasing practice.

Previous work with the anatomy of the ear program indicated that students who had not had previous course work in this area could not produce the correct anatomical names. The investigators therefore decided to dispense with the pretest and simply screened out those students who had completed previous course work in audiology or related areas. Three students were eliminated on this basis. The experimental Ss reported to the CAI Laboratory individually and were met by a proctor. Upon arriving at the Laboratory, each S was escorted to an experimental, sound-proofed room containing the typewriter terminal. The proctor then read the following instructions to the S:

> You see in front of you a two-way typewriter used for computer-assisted instruction. A computer presents problems via the typewriter, and the student enters his answers by typing them on the machine. We are engaged in an experiment to test the feasibility of teaching by computer. In order to get a fair evaluation, it is very important that you do the best job you can on the course which you will take.
>
> Following instruction, you will be given a test. You will also be required to come back for another phase of the study. This second session will require about 15 minutes of your time; however, you will not be paid unless you participate in both phases of the experiment.
>
> Before we begin the actual instruction, there are a few more things you may want to know. The time to complete the course varies depending upon the particular student. There are auxiliary materials you will need while working through the program. If your proctor has not specifically explained this to

you, please remind him to do so <u>now</u>. Be sure to refer to these materials at the appropriate time. As you work through the problems, you are not to look at previous problems in the course.

If you suspect any difficulty with the equipment, call on the proctor for assistance. However, the proctor has been instructed to give you no help with the problems or the content of the course. Do you have any questions before we begin?

Following the instructions, the proctor assisted the S in signing on to the appropriate experimental version of the course; either ordered sequence-one, two, or three trials, or scrambled sequence one, two, or three trials. As was the case in Experiment I, the scrambled sequence was again determined from a table of random numbers. The programming strategy employed in Experiment I (see Figure 5) was again employed in Experiment II to generate the six program versions from one basic ordered sequence program. At the time of sign-on, the S identified himself to the computer as a member of a particular treatment group, and the computer presented the appropriate experimental version of the program. In the two- and three-trial conditions of the scrambled sequence groups, the same scrambled sequence was repeated on the second and third trials rather than compiling a new scrambled sequence for each trial.

After the S had signed on to his appropriate treatment condition, the proctor remained with him through the first few frames of the program to insure that the S understood how to operate the terminal. Once assured that the S had adequately mastered the operation of the typewriter terminal, the proctor left the experimental room, and the S completed the instructional phase of the experiment. Immediately after the S had completed instruction, he was given a posttest (see Appendix C) to measure his ability to recall the names of the parts of the ear which were taught in the program. The Kuder-Richardson formula 20 reliability of the anatomy of the ear posttest was found to be .86 based on the sample of Ss in the ordered sequence condition. When the S had completed the posttest, he proceeded to complete the <u>Student Reaction Inventory</u> measure of attitude towards CAI. Following this phase of the experiment, each S was scheduled to return to the CAI Laboratory one month later to take a retention test to measure delayed recall. At this time, the <u>Student Reaction Inventory</u> was re-administered to obtain data on the stability of attitudes toward CAI over time.

In summary, the independent variables of Experiment II consisted of the course sequence (scrambled or ordered) and the number of practice trials through the program (one, two, or three cycles). The dependent variables consisted of errors made during instruction, instructional time, immediate recall as measured by the posttest, delayed recall as measured by the one-month retention test, and attitude towards CAI as measured by the Student Reaction Inventory.

One important addition to Experiment II was the inclusion of the one-month retention measure. The inclusion of a retention measure seemed most appropriate in view of the fact that the instructional program primarily involved the recall of anatomical labels. However; the retention measure was included for still another reason. Wodtke (1967) suggested that some instructional variations may have their primary effects on delayed recall rather than on immediate postinstruction performance. For example, one might predict that the effects of an inadequate instructional sequence could be overcome by a subject on an immediate posttest as a result of the S's own ability to reorganize and restructure the material, but that when the S was required to retain the material over a period of time, the scrambled sequence might interfere with recall. The fact that a scrambled sequence produced an increase in the number of errors during instruction in the two previous experiments, suggests that a larger number of incorrect responses might exist in the S's response repertoire at the time of measuring delayed recall. It is also possible that the logical structure of an ordered program aids delayed recall. The provision to include the delayed posttest in the present experiment provided some evidence on the effects of course sequence on retention over time.

Results

In general the results of Experiment II supported the experimenters expectations that the sequence of instruction in the anatomy of the ear program would not be as important a variable as in the modern mathematics program. The mean number of errors committed during instruction by the six treatment groups is shown in Table 10, and the analysis of variance of these data is shown in Table 11. All effects were nonsignificant statistically, although there was a slight tendency for the scrambled sequence groups to make more errors than the ordered sequence groups. Essentially the same results were obtained when instructional time was analyzed as the dependent variable. The mean total instructional time for the six treatment groups is shown in Table 12 and the analysis of variance of these data is shown in Table 13. The only clearly significant effect was that for cycles which is simply the result of the additional time required by the subjects in the three cycle

Table 10

Mean Number of Errors for the Scrambled and Ordered Sequence Groups Completing One, Two, or Three Cycles Through the Program

		Sequence	
		Ordered	Scrambled
Number of Cycles	1	$\bar{X}$ 8.70 s.d. 4.47 n=10	$\bar{X}$ 10.33 s.d. 4.81 n=9
	2	$\bar{X}$ 8.38 s.d. 3.31 n=8	$\bar{X}$ 8.73 s.d. 6.15 n=11
	3	$\bar{X}$ 10.10 s.d. 4.35 n=10	$\bar{X}$ 13.25 s.d. 7.36 n=12

Table 11

Analysis of Variance of Total Errors for Scrambled and Ordered Sequence Groups Completing One, Two, or Three Cycles Through the Program

Source	d.f.	M.S.	F	P
Sequence	1	47.56	1.46	$>.10$
Cycles	2	58.55	1.80	$>.10$
Sequence x Cycles	2	9.89	<1.0	
Error	54	32.43		

Table 12

Mean Total Instructional Time in Minute for the Scrambled and Ordered Sequence Groups Completing One, Two, and Three Cycles Through the Program

Number of Cycles	Sequence: Ordered	Sequence: Scrambled
1	$\bar{X}$=15 s.d. 8.31 (n=9)	$\bar{X}$=21 s.d. 9.34 (n=14)
2	$\bar{X}$=27 s.d. 8.86 (n=9)	$\bar{X}$=22 s.d. 8.85 (n=10)
3	$\bar{X}$=26 s.d. 8.70 (n=11)	$\bar{X}$=36 s.d. 8.83 (n=11)

Table 13

Analysis of Variance of Total Instructional Time, for Scrambled and Ordered Sequence Groups Completing One, Two, or Three Trails Through the Program

Source of Variation	d.f.	M.S.	F	P
Sequence	1	854,350	2.15	$> .10$
Cycles	2	3,331,950	8.40	$< .001$
Sequence x Trials	2	1,147,585	2.89	$< .10$
Error	58	396,517		

Note: The analysis in this table is based on seconds rather than minutes as in table 12.

condition to complete three cycles through the program. (The reader will recall that these effects were statistically significant in the modern mathematics program the P being less than .10 for instructional time, and less than .001 for errors.)

In an attempt to provide some additional insights into the effects of scrambled sequence upon errors committed during the course of instruction, a more detailed analysis of the error data was undertaken. The original analyses reported above were based upon a randomized groups design (i.e., separate groups of Ss were assigned at random to each of the six treatment combinations). If each S's error score is obtained separately for each cycle through the program, it is possible to analyze the error data for the three-cycle and two-cycle groups as a two-way factorial analysis of variance with repeated measurements on the cycles effect. Although the sample sizes of these analyses were somewhat smaller than in the original analysis (since they are based on only part of the original sample), the repeated measures analysis provided the added control for within subject variance. Table 14 and Figure 10 contain the results of the repeated measures analysis for the scrambled and ordered sequence groups (n = 10 per group) which were run in the three cycle condition. The sequencing main effect was nonsignificant statictically, the main effect due to cycles was significant at less than the .01 level, and the cycles by sequence interaction was significant at less than the .01 level. The significant cycles main effect reflects the improvement in performance as a function of repeated practice with the program. The significant cycles by sequence interaction suggests that students during the first cycle through the program were detrimentally effected by the scrambled sequence, but that by the time they had completed the second cycle, the sequencing effect was no longer evident. This finding is reminiscent of the finding in Experiment I, that the Ss in the scrambled sequence condition were able to overcome an initial increment in errors and eventually achieved a level of performance commensurate with that of the ordered sequence group. However, the evidence for an initial decrement in performance for the scrambled sequence group in Experiment II is much less consistent than was the case in Experiment I. For example, if one examines the repeated measures analysis for the subjects in the two-cycle condition shown in Table 15 and Figure 11 one now finds that the cycles by sequence interaction is nonsignificant. The only statistically significant effect in this analysis was that for trials.

Figures 12 and 13 show an even more detailed breakdown of performance reflected in terms of the mean number of errors per blocks of five instructional frames for each cycle through the program. The peaks in mean errors in frames 11-15 (block No.3) in the ordered sequence condition appear to reflect an increase in

Table 14

Repeated Measures Analysis of Variance of Errors Made in the Program for the Scrambled and Ordered Sequence Groups Completing Three Cycles Through the Program*

Source	S.S.	d.f.	M.S.	F	P
Between Ss	276.40	19	14.55		
A Sequence	29.40	1	29.40	2.14	
Ss Within Groups	247.00	18	13.72		
Within Ss	1415.33	40	35.38		
B Trials	850.63	2	425.32	35.95	< .01
AB	138.70	2	69.35	5.86	< .01
BX Only Within Groups	426.00	36	11.83		

*Note: In this an the other repeated measures analyses, some Ss were thrown out of some cells at random in order to achieve proportionality.

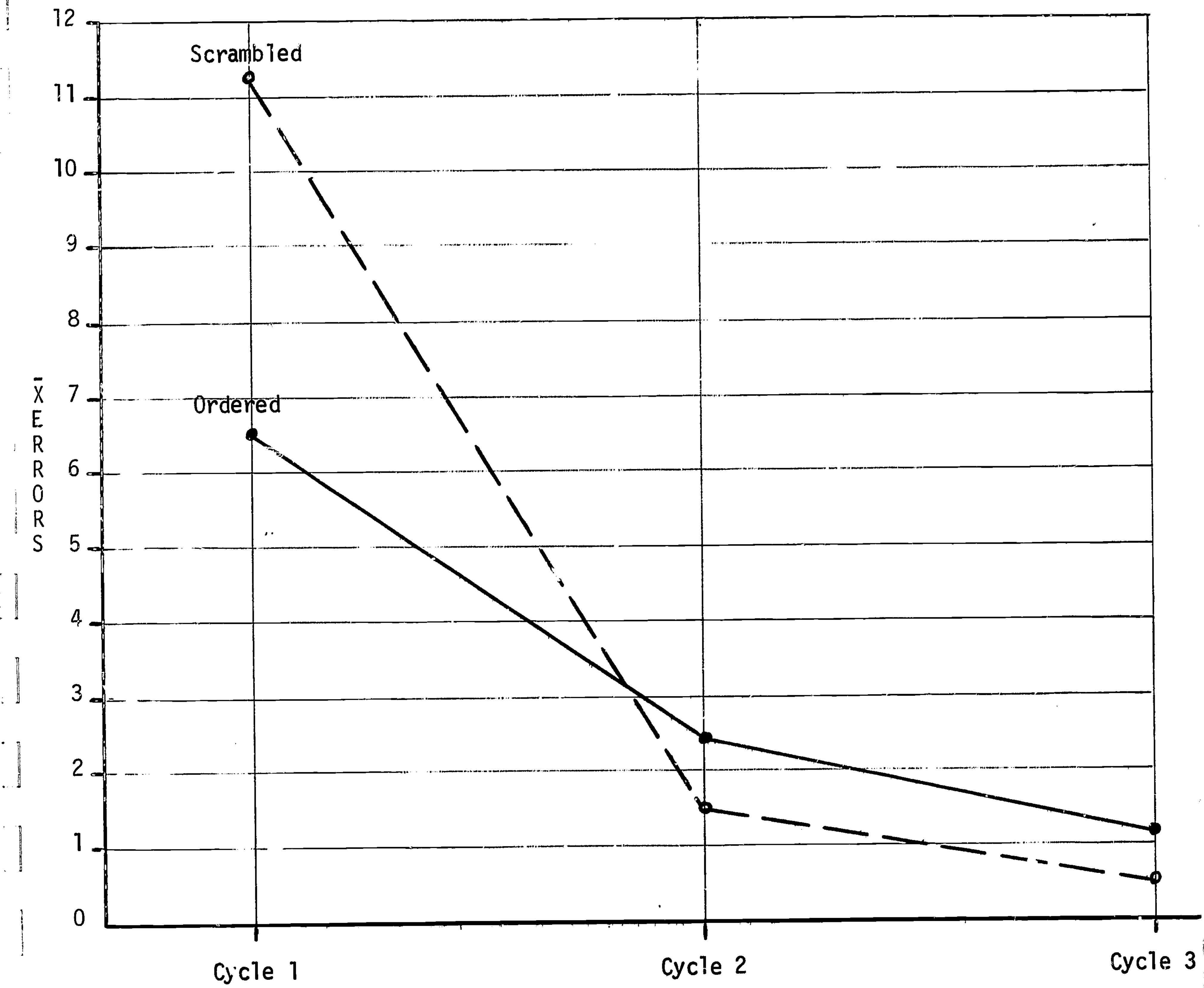


Fig. 10. Mean number of errors in program for each cycle through for the scrambled and ordered groups completing three cycles through the program. (N=10 for ordered group, N=12 for scrambled group.)

Table 15

Repeated Measures Analysis of Variance of Errors Made in the Program for the Scrambled and Ordered Sequence Groups Completing Two Cycles Through the Program

Source	S.S.	d.f.	M.S.	F	P
Between Ss	189.00	15	12.60		
A Sequence	.12	1	.12	< 1.0	
Ss Within Groups	188.88	14	13.49		
Within Ss	479.00	16	29.93		
B Trials	388.00	1	338.00	34.34	< .001
AB	3.13	1	3.13	< 1.0	
BX Ss Within Groups	137.87	14	9.84		

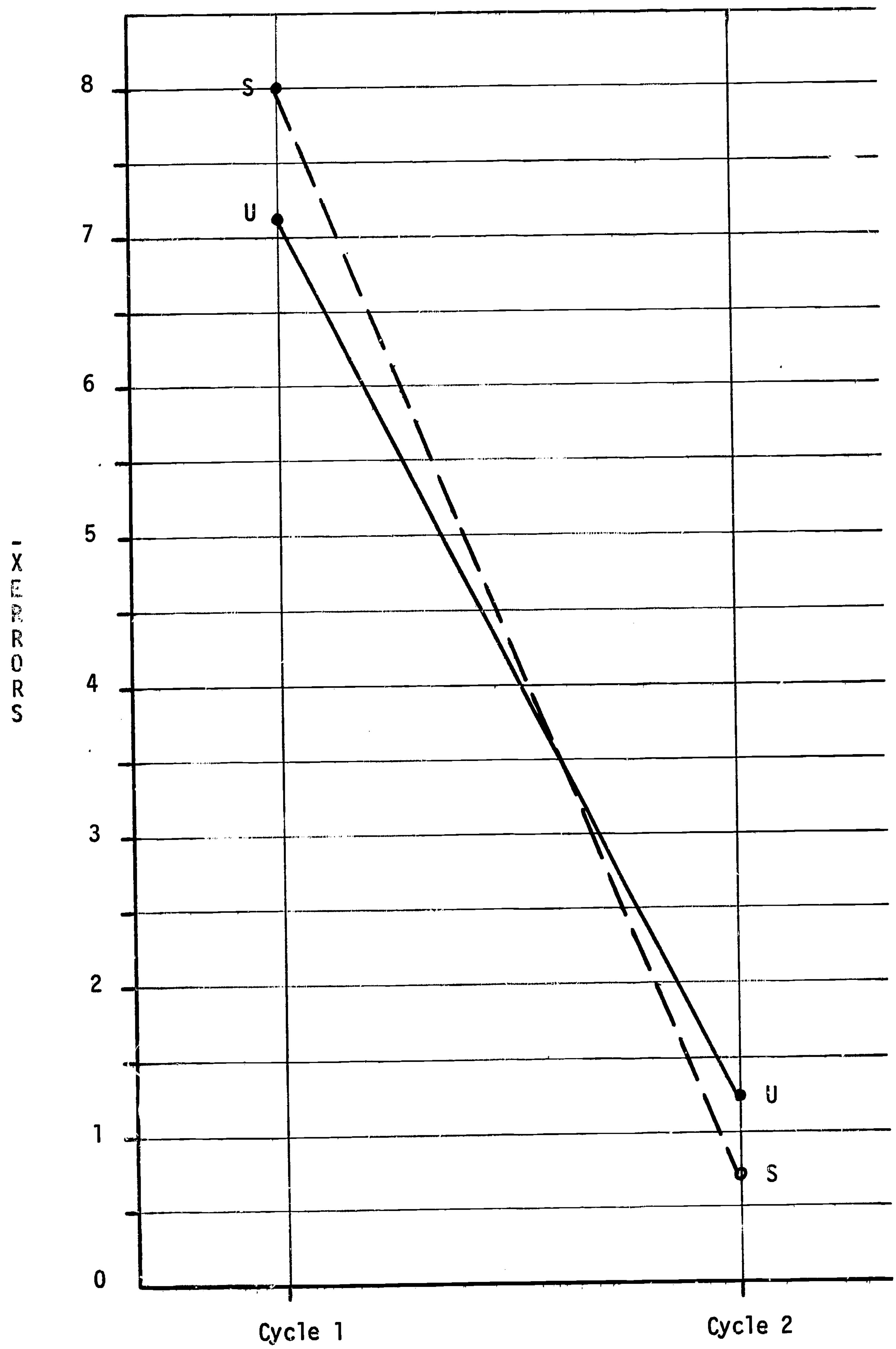


Fig. 11. Mean number of errors in program for each cycle through for the scrambled and ordered groups completing two cycles through the program. (N=8 for ordered group, N=11 for scrambled group.)

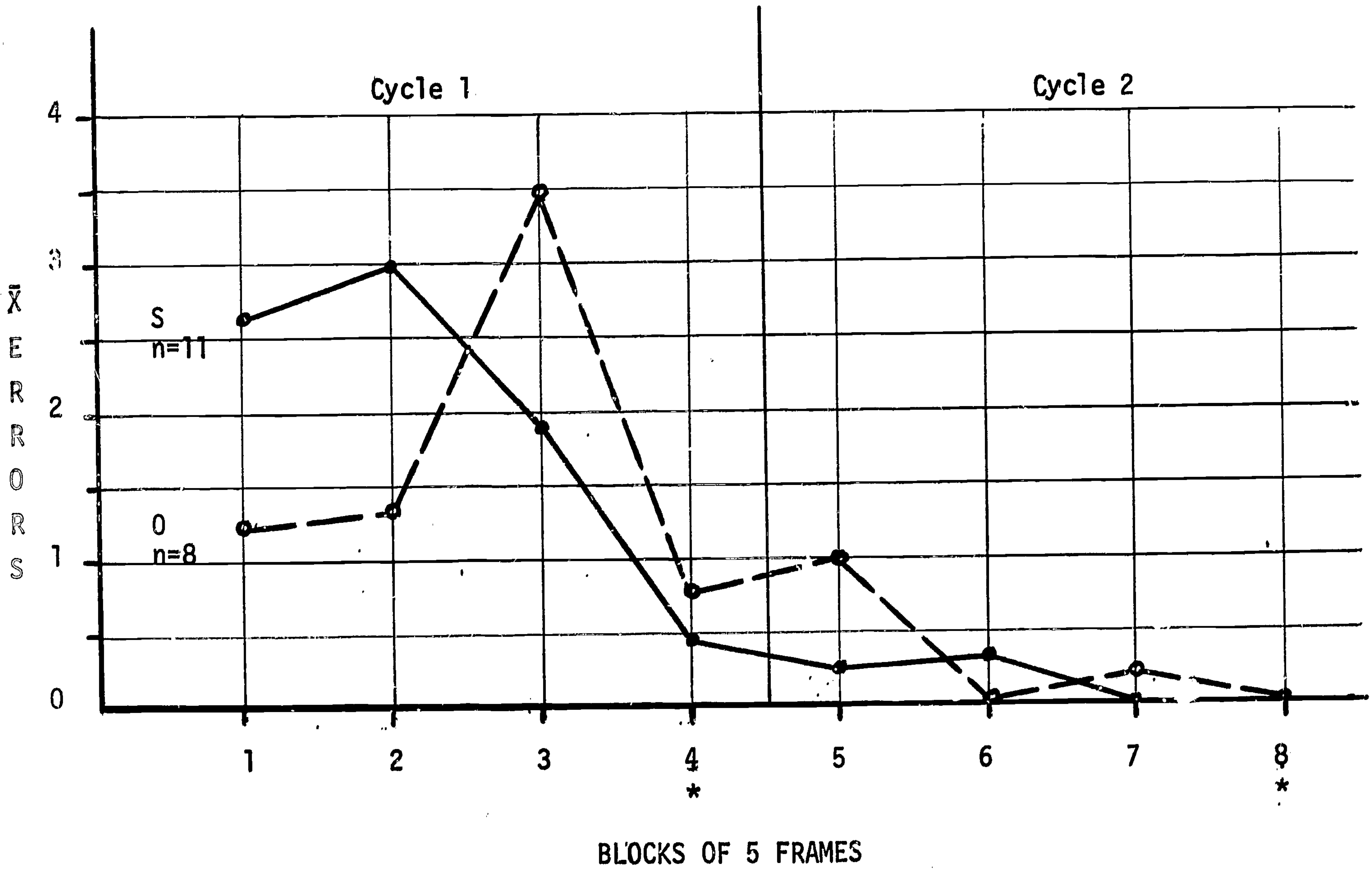


Fig. 12. Mean number of errors by blocks of five frames for the scrambled and ordered groups completing two cycles through the program

*Four Frames

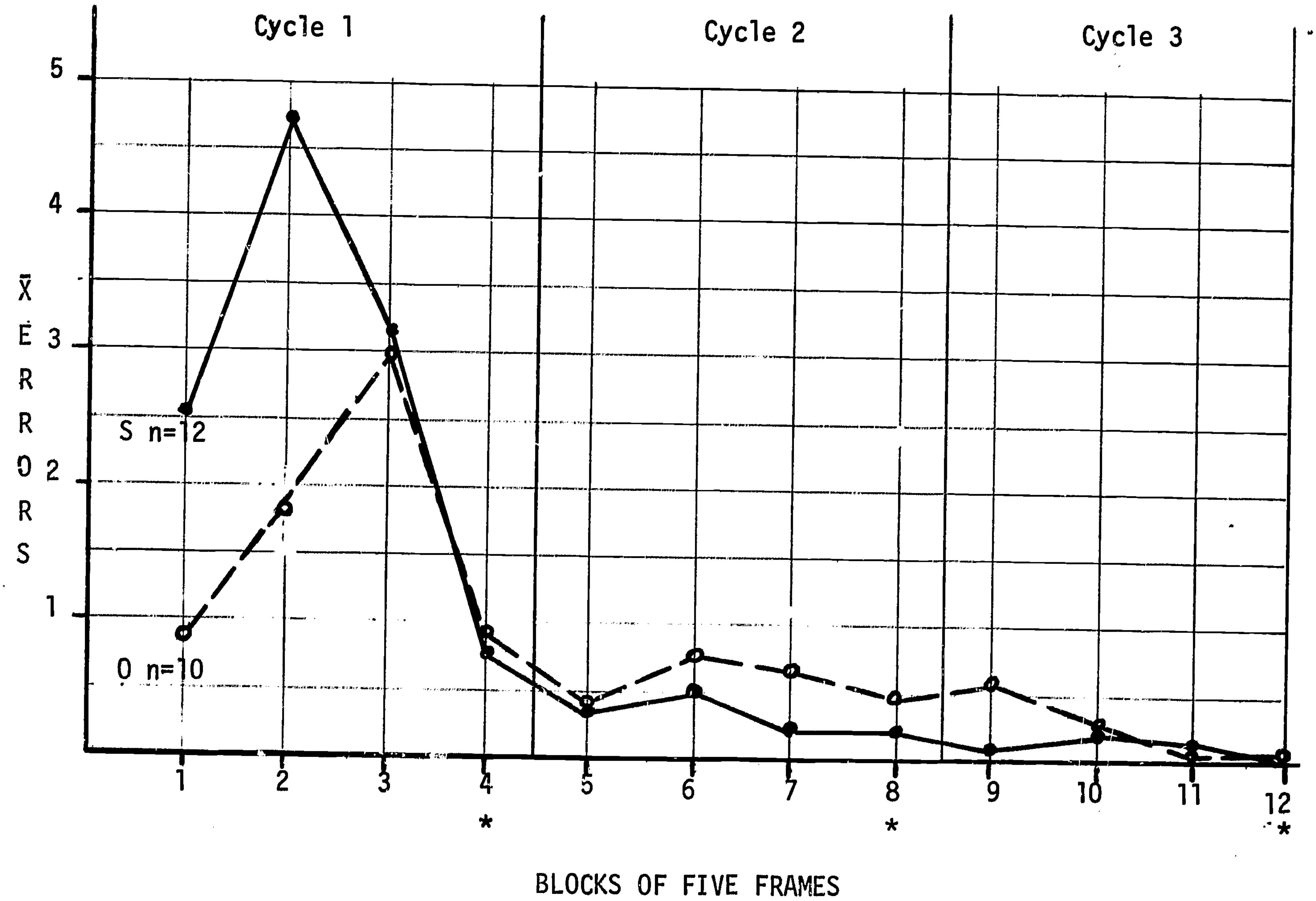


Fig. 13. Mean number of error by blocks of five frames for the scrambled and ordered groups completing three cycles through the program.

*Four Frames Here

the difficulty level of those frames. In each of the two figures, the scrambled sequence appears to have produced an increase in errors during the first ten frames, however, this effect is primarily due to the fact that the more difficult frames 11-15 in the ordered program were displaced in the scrambled program to the earlier segments 1-5 and 6-10, thus, accounting for the poorer performance in these segments for the scrambled sequence groups. The performance decrements of the scrambled sequence groups during the first ten frames do not appear to be the result of any disturbance of a "logical" sequence in the anatomy program, but rather resulted from a change in position of more difficult frames.

These results are taken by the writers to support the contention that performance on the anatomy program was not as dependent upon a logical sequence of instruction as was performance on the modern mathematics program. Scrambling the modern mathematics program produced highly significant increases in errors, while scrambling the anatomy program produced generally nonsignificant differences on the error measure. The most one can say concerning the effects of scrambling on the anatomy program, is that there may be the suggestion of a slight initial increase in errors during the early frames, which quickly dissipates as the student proceeds through the program. However, this effect was confounded in the present study with the difficulty of the frames in the early segment of the scrambled program.

Further evidence for the impotence of the sequence variable in the anatomy program comes from the analysis of the immediate posttest and one-month measures. These data were analyzed by means of a three-way factorial repeated measures analysis of variance. The three factors in the analysis were the number of practice cycles through the program (one, two, or three), scrambled or ordered sequence, and immediate versus delayed posttest (the single repeated measure). The means and standard deviations for the various treatment combinations are shown in Table 16. As can be seen from Table 16, the means within each of the tables for the immediate and delayed posttests are highly similar indicating a lack of an effect for cycles, sequence, or the cycles by sequence interaction. The means do show a marked, consistent drop from the immediate to the delayed posttest indicating considerable loss due to forgetting. The analysis of variance of these data shown in Table 17 simply bears out what has already been seen by examining the means in Table 16. The only statistically significant effect was that for the immediate versus delayed posttest which reflects the decrement in recall over time. The nonsignificant interaction between sequence and immediate versus delayed posttest indicates the lack of evidence for differential retention over time as a function of one or the other of the treatments.

Table 16

Means for the Immediate and Delayed Retention Tests for the Six Treatment Combinations

Immediate Posttest Sequence

Number of Cycles	Ordered	Scrambled
1	$\bar{X}$=12.1 s.d. 4.17 n=8	$\bar{X}$=13.5 s.d. 2.12 n=8
2	$\bar{X}$=12.1 s.d. 4.09 n=10	$\bar{X}$=13.1 s.d. 4.01 n=10
3	$\bar{X}$=15.6 s.d. 3 [illegible] n=11	$\bar{X}$=14.2 s.d. 4.88 n=11

One-month Retention Test

Number of Cycles	Ordered	Scrambled
1	$\bar{X}$=6.2 s.d. 2.28 n=8	$\bar{X}$=5.9 s.d. 4.70 n=8
2	$\bar{X}$=6.8 s.d. 3.84 n=10	$\bar{X}$=6.5 s.d. 3.04 n=10
3	$\bar{X}$=7.6 s.d. 4.48 n=11	$\bar{X}$=7.4 s.d. 4.67 n=11

Table 17

Analysis of Variance of the Immediate and Delayed Posttest for the Scrambled and Ordered Sequence Groups Experimenting One, Two, or Three Cycles Through the Program

Source	S.S.	d.f.	M.S.	F	P
Between Ss	1479.55	57			
Cycles	76.14	2	38.07	1.42	
Sequence	.14	1	.14	<1.0	
Sequence x Cycles	11.29	2	5.65	<1.0	
Error Between	1391.98	52	26.77		
Within Ss	1715.00	58			
Immediate vs. Delayed Retention	1311.21	1	1311.21	180.17	$<.001$
Cycles x Immediate vs. Delayed Retention	11.16	2	5.58	<1.0	
Sequence x Immediate vs. Delayed Retention	1.69	1	1.69	<1.0	
Sequence x Cycles x Retention	12.50	2	6.25	<1.0	
Error Within	378.44	52	7.28		
Total	3194.55	115			

The means, standard deviations, and analysis of variance of the Student Reaction Inventory measure of student attitude towards computer assisted instruction are shown in Tables 18 and 19. The effects of the treatments on attitude towards CAI were statistically nonsignificant.

Table 18

Means for Attitude Towards CAI for the Six Treatments Combinations

Attitude Immediately Following Instruction

		Sequence: Ordered	Sequence: Scrambled
Number of Cycles	1	$\bar{X}=78.0$ s.d. 9.38 n=6	$\bar{X}=72.0$ s.d. 11.35 n=9
	2	$\bar{X}=70.6$ s.d. 15.02 n=8	$\bar{X}=72.2$ s.d. 7.39 n=12
	3	$\bar{X}=77.8$ s.d. 9.79 n=8	$\bar{X}=71.0$ s.d. 10.64 n=12

Attitude One-Month Following Instruction

		Sequence: Ordered	Sequence: Scrambled
Number of Cycles	1	$\bar{X}=78.7$ s.d. 7.09 n=6	$\bar{X}=74.2$ s.d. 7.33 n=9
	2	$\bar{X}=66.4$ s.d. 13.60 n=8	$\bar{X}=71.7$ s.d. 16.91 n=12
	3	$\bar{X}=73.8$ s.d. 6.92 n=8	$\bar{X}=66.7$ s.d. 12.62 n=12

Table 19

Analysis of Variance of the Immediate and Delayed Attitude Measures for the Scrambled and Ordered Sequence Groups Experiencing One, Two, or Three Cycles Through the Program

Source	S.S.	d.f.	M.S.	F	P
Between Ss	13288.70	54			
Cycles	390.02	2	195.01	< 1.0	
Sequence	189.87	1	189.87	< 1.0	
Sequence x Cycles	580.56	2	290.28	1.17	< .10
Error Between	12128.25	49	247.52		
Within Ss	1955.50	55			
Immediate vs. Delayed Attitude	92.74	1	92.74	2.70	< .10
Cycles x Immediate vs. Delayed Attitude	144.89	2	72.44	2.11	< .10
Sequence x Immediate vs. Delayed Attitude	17.66	1	17.66	< 1.0	
Sequence x Cycles x Attitude	19.24	2	9.62	< 1.0	
Error Within	1680.97	49	34.31		
Total	15244.20	109			

Discussion and Conclusions

Overall effects of course sequence

The results of the present series of investigations are in general agreement with a number of previous findings (Roe, Case, and Roe, 1962; Levin and Baker, 1963) and in agreement with one recent study (Payne, Krathwohl, and Gordon, 1967) which all appear to indicate that the detrimental effects of scrambling a "logical" sequence of instruction are not as large as has been previously suggested in the educational folklore of subject matter sequencing. In general, the results of these previous investigations together with the results of the present series of studies indicate that criterion test performance is non-significantly different for groups given scrambled versus ordered programed instruction. However; the present results do not agree with two previous studies (in which error and instructional time measures were reported), with regard to the effects of item scrambling on errors committed by the Ss during instruction and instructional time. Roe, Case, and Roe (1962) compared a scrambled versus ordered instruc ional program in elementary probability using college student Ss and found nonsignificant differences on errors made during instruction and instructional time. Levin and Baker (1963) using an instructional program in informal geometry with second grade children obtained similar nonsignificant effects for error and instructional time scores. These findings contrast with the results of the present pilot study and Experiment I in which scrambling the sequence of instruction in a modern mathematics program was found to significantly increase errors and instructional time. Thus, our results indicated that scrambling the instructional sequence did decrease the efficiency of instruction for the modern mathematics program as measured by the amount learned per unit of time, but that scrambling did not effect the final level of criterion performance achieved by the Ss. Eventually Ss in both the scrambled and ordered sequence groups achieved approximately the same level of criterion test performance, with the scrambled group taking longer to attain that level and making more errors during instruction than the ordered group.

These results were frankly a surprise to the investigators. Motivation for the project stemmed from a skepticism of some of the earlier findings. We frankly expected to obtain large and unambiguous sequence effects in the modern mathematics program and nonsignificant sequence effects in the anatomy of the ear program. Although the latter expectation was generally confirmed, the effects for the mathematics program were much smaller than had been expected. For example, Table 4 shows that in Experiment I, the instructional time variable was significant at only the 10 per cent level of statistical significance, and that the scrambled sequence Ss averaged only nine minutes longer to complete the two-hour program than the ordered sequence group.

There are several possible explanations for these results. The investigators believe that Experiment I avoided several ambiguities of interpretation which were present in earlier studies. It was possible to conclude that the students employed in the study were almost totally naive with regard to the conversion of numbers to a nondecimal base. The pretest performance of the Ss indicated that they had literally no idea of what was required for solution of the problems prior to instruction. This result contrasts with earlier reports in which naivete with regard to the subject matter was either assumed, or when Ss achieved a high level of pretest performance prior to instruction indicating a high degree of prior knowledge of the subject matter (see for example the recent study by Payne, Krathwohl, and Gordon, 1967, in which a group which did not receive instruction obtained a relatively high mean score on a measure of achievement of elementary concepts in psychological measurement). If the Ss employed in an experiment already know a substantial amount of the subject matter to be taught, it is extremely difficult for the experiment to detect any treatment effects, let alone the effects of instructional sequencing. Experiment I minimized the effects of Ss' prior knowledge of the subject matter. With the exception of one or two Ss who achieved one or two points on the pretest, all Ss obtained pretest scores of zero on the ability to convert numbers from one number base to another.

Another problem which has arisen in previous studies of sequencing concerns the effectiveness of the instructional program employed in the experiment. If the "ordered" sequence program is only modestly effective in producing mastery of the subject matter, it is unreasonable to expect very large differences between that program and a scrambled sequence version. In addition to the problem of the Ss prior knowledge of the material, the failure of the instructional program to produce mastery also decreases the sensitivity of the experiment to the treatment effect. Figure 4 on page 22 indicates that the number systems program employed in Experiment I produced a high degree of mastery. The reader should recall that the Ss whose criterion performance is shown in Figure 4 all obtained scores of zero on the pretest of this ability. Thus, the program achieved its objectives for the large majority of the Ss. The lack of a substantial treatment effect in Experiment I cannot be attributed to a failure of the instructional program to teach the intended material. The writers believe that the selection of naivé Ss, and the effectiveness of the instructional program maximized the sensitivity of the experiment to the possible effects of scrambled sequencing (assuming of course that such an effect exists for the subject matter and population used).

What then could have accounted for the lack of a sequence main effect (aside from the error score) in Experiment I? One possibility, which has created difficulties in the interpretation of previous findings, is that the a priori judgment of the investigators concerning the inherent hierarchical sequence of the modern mathematics program was incorrect. If a subject matter did not have an inherent "order" or conceptual hierarchy, then one would not expect the sequence of instruction to be a crucial variable. In fact, this was precisely the rationale for Experiment II which employed a topic seemingly devoid of a conceptual hierarchy. The lack of a conceptual hierarchy within the subject matters may account for the absence of sequence effects in the three previous experiments discussed above (Roe, Case, and Roe, 1962; Levin and Baker, 1963; and Payne, Krathwohl, and Gordon, 1967). Each of these studies failed to find statistically significant differences in the number of errors committed during instruction for groups given a scrambled as compared to an ordered sequence. It is difficult to believe that scrambling the sequence of an inherently hierarchical subject matter would not increase the frequency of errors made during instruction. To return to an earlier example, if a subject matter has the inherent conceptual sequence A, B, C, and D, and one teaches concepts B, C, or D first in the program it is literally impossible for a subject to solve problems correctly in these segments without the prerequisite training (assuming the hierarchical model is correct for the given subject matter). Measures of program performance such as errors and instructional time are the only independent empirical tests one has for determining the validity of an assumed, a priori conceptual hierarchy. If within program performance measures are not affected by altering the sequence of a presumed "logically" ordered subject matter, then one is left with the inescapable conclusion that the original hierarchical model of the sequential characteristics of the subject matter was incorrect. The writers take the lack of significant differences on within program error measures in these earlier studies as strongly suggestive that the programs employed were not in fact "logically" ordered, at least not in the sense of the conceptual hierarchy model suggested by Gagné (1962) which implies that mastery of subordinate tasks in the hierarchy is essential to the mastery of tasks at higher levels in the hierarchy. Destroying the sequence implied by the hierarchical model for any given subject matter would have to depress performance on higher level tasks assuming the model was valid for that subject matter.

The results of these previous studies may be contrasted with the results of Experiment I. The fact that the scrambled version of the modern mathematics program did produce a significant increase in within-program errors strongly suggests the presence of a facilitative sequence in the ordered version of the program. Performance on later program segments was facilitated by having had the necessary instruction on earlier prerequisite segments. The skeptical reader is again referred to Figure 6 page 29. The dashed performance curve shows the performance of the scrambled sequence Ss on blocks of ten scrambled frames. The mean number of errors starts much higher than that for the ordered sequence group and does not drop below the ordered group until the last twenty frames of the program. The broken dash-dot performance curve shows the increase in errors resulting from the scrambling of the original blocks of ten frames. The writers take these data to support their contention that the ordered version of the modern mathematics program employed in Experiment I did contain an ordered conceptual sequence. We would conclude that the lack of a detrimental effect for scrambled sequencing on the criterion test therefore, did not result from a lack of conceptual sequence in the subject matter program, and must be accounted for through some other explanation.

Several explanations have been offered by researchers to account for the apparently small effects of scrambling the sequence of an instructional program on student performance. Roe, Case, and Roe (1962) suggested that scrambling the sequence of instruction increased student motivation to master the task, and consequently increased the strength of reinforcement when the student finally did achieve insight into the material. These authors state that: "Presenting items out of sequence possibly introduced a task oriented anxiety which was subsequently relieved in a moment of revelation when a missing clue was discovered." (Roe, Case, and Roe, 1962, p. 104) The only evidence in the present studies bearing on this hypothesis were the responses of students to the Student Reaction Inventory in the pilot study. Students in the scrambled sequence condition reported being significantly more tense than students in the ordered program group. These student self-reports did support the contention of Roe, Case, and Roe that item scrambling increases anxiety, however, whether increased student tension actually facilitates learning is a debatable question which will require further research. In fact, contrary to the notion of Roe, Case, and Roe, previous research on the effects of anxiety on learning strongly suggests that high anxiety is detrimental to the learning of complex tasks. Classroom experience in teaching quantitative courses such as statistics, and measurement also suggests that high anxiety in the students can interfere with learning.

Payne, Krathwohl, and Gordon (1967) suggested another interpretation which relies heavily on the organizational powers of the learner. These investigators state that the scrambled sequence program,

> "...was viewed as a large step program, the individual sorting out items that were previously contiguous or nearly so and recalling from memory the items that relate to a currently attended problem as appropriate. Incorrectly solved items were reviewed in relation to the correct answer, to determine the correct principle, relation or fact. In many instances this resulted in a discovery and inductive development of information and principle that permitted the students to learn the material involved at an earlier point (with fewer steps) than they would have gone through in the logically sequenced program. Thus the learner inductively built the knowledge and principles as he went through the program, in many instances using almost a discovery method of learning." (Payne, Krathwohl, and Gordon, 1967, pp. 131-132)

The results of Experiment I appear to support this interpretation. Figure 6 on page 29 shows the within program performance (as measured by mean number of errors per block of ten frames) of the scrambled and ordered sequence groups. The mean number of errors for the scrambled sequence group started considerably higher than that for the ordered sequence group, and remained higher through the first forty frames of the program. Lacking the information necessary for the solution of the problems they encountered during the early segments of the program, the scrambled sequence Ss naturally had greater difficulty with the material and made more errors. However, the striking trend in the performance of the scrambled sequence group shown in Figure 6 is towards a decrease in errors, until by the last twenty frames in the program they were making fewer errors than the ordered sequence group. In spite of the scrambled sequence and the initially high error rate, these Ss were apparently able to reorganize the material and eventually achieved mastery of the concepts of the program. This interpretation is further supported by the finding that when the scrambled sequence Ss were given the criterion test immediately following instruction they demonstrated a level of achievement comparable to that of the ordered sequence group. It is important to note that in the present study, the Ss were able to achieve understanding of the scrambled material even though they were prevented from looking back at previous frames encountered in the program. The Ss would have to have recalled earlier material which was relevant to the solution of later problems.

The ability to reorganize scrambled material is undoubtedly a function of the cognitive development of the learner. Although it appears likely that college students are able to accomplish such reorganization, the writers would be extremely reluctant to generalize such a conclusion to the problem of sequencing learning materials for young children. It may be that sequencing is much more crucial in the education of young children who have not yet developed their own learning strategies.

Although the writers do not view the present results as justification for ignoring course sequencing in the preparation of instructional materials, we do believe that the present results together with those of a number of previous investigations raise some embarassing questions for an approach to instruction which places a strong emphasis on optimal course sequencing, small step programs, minimal error rates, etc. Our results suggest that students are highly adaptable to a scrambled sequence of instruction and are able to bring their own organization (and perhaps more meaningful organization for them) to bear on the subject matter. Although the effect of sequence does seem to depend on the nature of the subject matter (as Experiment II indicates), and on the level of cognitive development and abilities of the learners (as the interaction data in Experiment I suggests), we would venture the tentative conclusion on the basis of the accumulated research to date, that instructional sequencing may be an overrated variable. Much research in the area of systematic instruction has focused on what may be relatively trivial manipulations of the instructional stimulus (e.g., sequence, length of frame, size of step, overt versus covert response, feedback, etc.). Such manipulations of instructional stimulus variables may account for only a very small portion of the variance in student learning, compared to the amount of variance accounted for by the learning strategies and information processing strategies employed by students. A purely stimulus-oriented approach to the development of instructional materials may produce only very limited gains, if the most important componant in the instructional process is the student with all of his aptitudes, interests, and his ability to discover the concepts of a subject matter regardless of minor variations in the nature of the instructional stimulus or its presentation sequence. The writers were surprised in the present study by the ability of college students to achieve mastery of the material in spite of what appeared to be a most serious distortion of the sequence of the instructional material. Studies of the information processing strategies employed by learners in dealing with different variations in the instructional materials might provide more insights into how one might optimize the learning experience than comparisons of arbitrary manipulations in the instructional stimulus per se.

If one examines the genesis of the principles of programed instruction such as small-step, minimum error rates, ordered sequence, etc., one finds that most of these notions stem from research by Skinner and his colleagues on the learning of a very low level of animal life (e.g., the pigeon). This research led to a series of empirically based S-R laws which essentially ignored the subject or "O" variables. Perhaps with an animal which is as stimulus bound as the pigeon, one can afford to ignore the "O" variables, however, one runs great risks in generalizing laws derived on pigeons to complex subject matter learning with human beings. As one prominant psychobiologist recently put it in commenting on the dangers of generalizing research findings from pigeon to man, "The pigeon doesn't even have a cortex." Obviously the college students employed in the present study had a cortex, and our results suggested that they were able to put their cortex to good use in making sense out of a morass of instructional inputs. Ignoring such learner variables in research on instruction may lead to misleading conclusions concerning the importance of certain stimulus variables, and may hamper progress towards an individualized instruction which should take into account the capacities of the learner for processing information.

Interactions between student individual differences and instructional sequence

Several previous investigations have studied the problem of the interaction between student individual differences in the form of aptitude measures, and the effect of a scrambled sequence of instruction. Levin and Baker examined the intercorrelations of IQ measures with their dependent variables within each of the two treatment conditions, and found no consistent differences among the correlations (i.e., no evidence for an IQ by sequence interaction). Payne, Krathwohl, and Gordon, (1967) also looked for consistent trends in the correlations between the MSU Arithmetic Test and criterion performance within the scrambled and ordered sequence conditions. Again no consistent differences in the correlations for the two sequence conditions were evident, providing no evidence for an interaction effect between arithmetic test performance and sequencing. One of the most provocative findings reported on this problem were obtained by Stolurow (1964). Stolurow employed a "mixed" sequence and a "consecutive" sequence for teaching fractions to educationally handicapped high school students. (Mean mental age was equal to 12.25 years.) In the mixed sequence, the students could not determine which fraction would come next. In the consecutive sequence, the student, assuming he had learned the natural order of numbers, could anticipate the next in the series of problems. Full scale IQ, and total language scores correlated .61 and .63 respectively with posttest scores for students given the mixed sequence program, but did not correlate significantly with performance on the consecutive sequence program! Stolurow (1964) interprets these results as indicating that, "...the best sequence did for the poorest ability group what the highest ability groups could do for themselves regardless of sequence." (Stolurow, 1964, p. 351) In terms of the rationale of the present series of investigations, the high ability students were presumably able to reorganize the material on their own, whereas the low ability students were unable to reorganize on their own and had to rely upon the organization built into the instructional program. It is important to note in considering Stolurow's results in relation to the results of Levin and Baker (1963), Payne, Krathwohl, and Gordon (1967), and the results of the present investigation, that Stolurow's students were educationally mentally handicapped Ss. It is entirely possible that the aptitude by sequencing interaction exists when one is dealing with students of extremely low ability, but that the effect is nonexistent when one is dealing with a more representative sample of ability, or with a more able group of college students.

The results of the present series of investigations, although they do not present an entirely consistent picture, do provide some evidence which is supportive of the original hypothesis concerning the aptitude by sequencing interaction. The unexpected direction of the interactions obtained in the pilot study (i.e., high aptitude Ss seemed to be most detrimentally affected by the scrambled sequence at borderline significance levels) were clearly not replicated in Experiment I. In fact, the results of Experiment I conform to the results of Stolurow and support the notion that low aptitude Ss are most detrimentally affected by scrambling the sequence of instruction. The methodological improvements of Experiment I as compared to the problems encountered in the pilot study lead the writers to place more faith in the findings of Experiment I.

The unexpected interaction which appeared in the pilot experiment may have resulted from the extreme novelty of CAI and certain characteristics of the unrevised modern mathematics program which was employed in the pilot experiment. When the pilot study was conducted, CAI was very new on the Penn State campus, and for that matter very new throughout the country. Few of the students employed in the pilot study had the slightest idea at that time as to what CAI was all about. The subjective impressions of the investigators while running the subjects for the pilot study, was that the students evidenced a high degree of ego involvement and a desire to do well at this highly novel method of instruction. Anxiety generated by the extreme novelty of computerized instruction, and anxiety stimulated by the difficulty encountered by the high aptitude Ss (in view of the usual expectations of success of such groups), may have resulted in the unusual decrements in the performance of the high ability Ss. In addition, the scrambled program employed in the pilot study seemed to produce much more anxiety in the Ss than the revised program employed in Experiment I. This anxiety probably resulted from the fact that the modern mathematics program employed in the pilot experiment was much more of a continuous verbal program than the revised program employed in Experiment I. Scrambling the original program, in addition to scrambling the conceptual sequence, also scrambled the continuous verbal flow of the program which seemed to cause a great deal of frustration and anxiety in the subjects. Subjects in the scrambled group in the pilot investigation took on the average 45 minutes longer to complete instruction than the Ss in the ordered sequence group. This time decrement may be compared with the much shorter 9 minute decrement experienced by the scrambled sequence group in Experiment I. It was not unusual while running subjects for the pilot investigation, to have a student bound out of the experimental room and indicate extreme frustration with his inability to figure out what the program wanted. Our impressions of the frustration and anxiety of the Ss in the pilot experiment were born out by the self-report responses to the Student Reaction Inventory which indicated a high level of tension in the scrambled sequence-high aptitude groups. The effects of frustration which were observed during the pilot study, simply were not observed during Experiment I The revised modern mathematics program, unlike the earlier unrevised program, consisted of a set of frames which essentially stood alone as separate problems in the conceptual sequence. The revised program contained a minimum of sequential dependencies based on the written continuity of the program. The disruption of the sequence of a verbally continuous program may cause learning difficulties which are not encountered when a program consists of frames which stand alone, each providing a separate problem to be solved. In addition, by the time Experiment I was being conducted, CAI had been on the Penn State campus for several years. It is highly likely that students from the subject population used had seen video tapes and other demonstrations of CAI, and the possibility of anxiety resulting from an extremely novel form of computerized instruction may have been reduced. The unusual interactions between aptitude and sequencing which were obtained in the pilot study may have resulted from high anxiety generated from novelty of the experiment, and the extremely frustrating characteristics of the original unrevised program.

The more reliable data of Experiment I, the improved instructional program, and the complete random assignment of Ss to treatments leads the writers to the tentative acceptance of the original directional prediction concerning the nature of the aptitude by sequence interaction. Although the two methods of analysis (correlational and the factorial analysis of variance) were not equivalent in statistical significance, these analyses produced results which were consistent in direction and statistically significant in the case of the factorial analysis of variance. This difference in the sensitivity of the two methods of analysis may have resulted from the greater power of the factorial analysis of variance for detecting true interaction effects. We take the results of Experiment I as tentative support for the conclusion that the effects of scrambling an instructional program containing a logical conceptual hierarchy do in fact depend to some extent on the aptitudes of the learner. The results suggested that the high aptitude Ss were able to reorganize a scrambled presentation of subject matter to an extent superior to that of the lower aptitude Ss whose performance was more detrimentally affected by item scrambling. However, here again the largest and statistically significant effects were for the error and time scores. The effect for criterion test performance was of smaller, statistically nonsignificant magnitude.

The differences in probability values obtained in the correlational analysis of the aptitude by sequence interaction effect, and the factorial analysis of variance approach raise some very interesting questions concerning the comparative power of these two procedures, and their appropriateness for examining interactions between individual difference measures and instructional treatments. Methodological questions necessarily arise when two methods of analysis applied to the same data produce discrepant P-values. Cronbach (1957) has advocated the comparisons of within groups regressions as a method of examining interactions among individual difference measures and experimental treatments. On the other hand, the more traditional approach to the problem has been to dichotomize high and low groups on the individual difference measure and run a factorial analysis. If the two approaches do not produce identical results, then the question arises as to why they do not, and which approach is the most appropriate or sensitive method for detecting interactions. Since such interactions are often of primary importance in educational research, this is a methodological question with important practical implications. As an outgrowth of the present project, the two senior authors are presently conducting a Monte Carlo experiment to determine the power of these two statistical methods for testing hypotheses concerning interaction effects. It is hoped that this investigation will provide some answer to the question of which method is most appropriate for studying interactions between individual difference measures and instructional treatments.

Subject matter characteristics and the effects of scrambled sequence

One of the original contentions of the present writers was that some of the inconsistency in the findings of previous studies of item scrambling might be due to differences in the subject matters employed. It seemed likely that the effects of scrambling the sequence of an instructional program would depend on whether the subject matter contained sequential dependencies among the concepts to be taught. Thus, if a subject matter consisted of a hierarchy of concepts (Gagné, 1962), and the hierarchical model had been validated for that subject matter, then scrambling the sequence of instruction for such a subject matter would _by definition_ affect _within program_ performance. If a subject matter consisted of a relatively unrelated set of facts, then one would not expect the sequence of instruction to be a critical variable. Previous research with the number systems and anatomy programs employed in Experiments I and II suggested that these subject matters were widely separated on the continuum of conceptual hierarchy (number systems program) to a subject matter containing a set of discrete facts (anatomy of the ear program). These two subject matters were chosen for the investigations in an attempt to maximize the difference between the programs on the degree of conceptual hierarchy inherent in the content. As indicated above, the investigators expected the scrambled sequence to have large detrimental effects on student learning of the hierarchial program, but little or no effect on the learning of the set of relatively discrete facts. The expectations with regard to the anatomy of the ear program were confirmed, however, the differences obtained for the number systems program were much smaller than had been expected. Although scrambling the sequence of the number systems program did effect within program performance (increased errors and instructional time), supporting our a _priori_ notion that the number systems program did contain a conceptual hierarchy, by the completion of instruction students had achieved mastery of the content as measured by an immediate criterion test in spite of the scrambled sequence.

The results of Experiment II with the anatomy program indicated almost consistently nonsignificant differences between the scrambled and ordered sequence groups on errors, instructional time, immediate retention, delayed retention, and attitude towards CAI. The only two clearly significant effects were those for cycles through the program and retention over time reflecting the failure of recall on the delayed retention measure. These data tend to support the conclusion that item sequence is less important for a program of discrete facts than for a program containing a conceptual hierarchy.

One qualification on the findings of Experiment II concerns the matter of the comparative length of the two programs employed. To attribute the lack of differences in Experiment II to a lack of a conceptual hierarchy in the anatomy program, it would have been desirable to have a program which was equivalent in length to the number systems program employed in Experiment I. In effect, Experiment II confounded program length with the conceptual characteristics of the subject matter. Unfortunately, it was simply not possible in the present study to utilize an anatomy program which was two hours in length. Preliminary experimentation with a much longer anatomy-of-the-ear program (two to three hours in length) indicated that students could recall very few of the anatomical terms after only one cycle through the program. A two-hour version of the anatomy program is highly analogous to a paired associates nonsense syllable list which would take Ss two hours to complete one trial through the list. Learning would have been so poor after only one two-hour cycle through the program that it would have been impossible to detect any differences attributable to the sequence of presentation. As the reader can see in Figure 10, even for the shortened version of the anatomy program, one cycle through the program was insufficient to produce adequate learning as indicated by the relatively high error rates during the first cycle through the program. This means that to use a two-hour program, it would have been necessary to cycle students through the program a number of times in order to obtain an adequate level of learning. Assuming that three cycles would have been sufficient, this would have necessitated approximately six hours of instruction (not to mention six hours of computer time per student), in addition to approximately one to two hours of criterion testing. It was not possible to expect such large time commitments from the students available for the investigation. Even if it had been possible to use the two-hour version of the anatomy program it is not at all clear to the writers in what sense the "length" of the number systems and anatomy programs would have been equated. Equating the length of a program involving rote learning with a program involving the learning of meaningful principles produces differences in the level of learning achieved. Thus, as indicated above, one cycle through the two-hour anatomy program would have been approximately equal in length and number of frames to the number sytems program, however, the level of learning would have been far lower than that obtained with one cycle through the two-hour number systems program in which the meaningful nature of the content facilitated learning. The question is do you equate program length or the level of learning achieved through the programs? Our decision was to attempt to equate the programs, to the extent possible, in terms of the level of mastery achieved by Ss following instruction in the ordered versions. We chose a shorter anatomy program, (necessitated in part by the demands on computer time usage and student time), and varied the number of practice cycles through the program from one to three in order to obtain a program which produced an adequate level of learning within the time limitations available. If we had held instructional time constant for the two programs we would not have had comparable amounts of learning in the two programs, and the results would have been confounded with differences in the levels of learning achieved in the two programs. We finally decided that it was more important to compare two programs which produced comparable and relatively high levels of learning, than it was to hold the length of instruction constant which would have resulted in inadequate levels of learning in the anatomy program.

In spite of the above difficulties encountered in equating the length of instruction for the two subject matters (number systems and anatomy of the ear) the writers are fairly confident that the results obtained with the shortened version of the anatomy program would be predictive of results obtained with a longer version of that program. Considerable experience with this program suggests that the anatomy subject matter is almost totally devoid of a conceptual hierarchy. Although lengthening the program would have increased the number of trials needed to reach an adequate level of recall, it would not be expected to alter the present results which indicated a lack of effect due to scrambling the sequence of instruction.

Computer-assisted instruction as a vehicle for research on instruction

To the writers' knowledge, the present project is one of the first attempts to conduct research on variables related to the instructional process within the context of computer-assisted instruction. It would probably be useful to researchers who are anticipating doing similar research in the CAI context to have the reactions of the present investigators to CAI as a research tool. The writers are in general agreement that CAI presents a potentially powerful research tool for investigations of the instructional process. In fact, we are inclined to the view that CAI may be more useful as a research tool than as an instructional device. There are several more obvious advantages to using CAI in educational research. CAI enables the researcher to have much more control over the presentation of the instructional treatments, and consequently can provide a much more standardized instructional treatment than has been possible in instructional research in the schools. Using the decision logic and capabilities for rapid revision of CAI it is possible for the researcher to produce almost instantaneous variations in the instructional treatment. For example, in Experiments I and II we used the decision logic of the computer to produce the scrambled and ordered versions of the program from one basic set of programed materials. In addition to these advantages, the computer also provided record keeping capabilities which allow an investigator easy access to information on student performance such as error and response latency measures.

Although there are many advantages to CAI research, the investigator just starting out in this area should probably be prepared to encounter some difficulties during the early phases of his research. Any new device as complicated as a CAI instructional system requires a rather lengthy shake-down period. This would be especially true for the first generation of CAI systems. The second generation systems (for example the IBM 1500) have incorporated many improvements based on the experiences of researchers with first generation systems. Some of the problems which have been encountered during our own research with a first generation CAI system have been interruption of Ss during the experimental instruction resulting from machine failure, loss of student records in storage resulting from a machine failure, difficulties encountered in attempting to retrieve student performance data from computer storage for experimental analysis, and errors in compiles of experimental instructional programs introducing systematic and unsystematic variations into the instructional treatment. One great aid to CAI research would be improved programs to retrieve student performance data in a form appropriate for analysis by the investigator with a minimum of turn around time. Improvements of the type required to make CAI a more useful research tool will undoubtedly occur as more experience is gained in the CAI field.

Summary

The present series of investigations was undertaken to clarify the findings obtained in several previous studies which suggested that scrambling a "logical" sequence of instruction did not appear to have a significant detrimental effect on learning (see for example Roe, Case, and Roe, 1962; Levin and Baker, 1963; and Payne, Krathwohl, and Gordon, 1967). These results ran somewhat counter to the educational folklore concerning the importance of instructional sequencing in subject matter learning. The present investigators posited that the effects of scrambling the sequence of instruction probably depended to some extent on characteristics of the subject matter being taught and individual differences among the learners. If a subject matter contained a sequential hierarchy of concepts (such as that suggested by Gagné, 1962) then one would expect scrambling the sequence to be detrimental to learning. If on the other hand, the subject matter consisted of a relatively discrete set of facts, one would not expect the sequence of instruction to be such a critical variable. One pilot investigation, and two complete experiments were undertaken to test the effects of scrambled versus ordered course sequencing with two subject matters. One subject matter in modern mathematics was selected because it appeared to contain a hierarchy of concepts. The second subject matter on the anatomy of the ear was selected because it seemed to represent the other extreme consisting primarily of a set of relatively unrelated facts. It was expected that the scrambled instructional sequence would be highly detrimental to the learning of the modern mathematics program, but have little or no effect on learning in the anatomy program. In addition to the subject matter considerations, interactions between the aptitude of the learners as measured by the Scholastic Aptitude Test (College student Ss were employed) and instructional sequence were examined to test the prediction that a scrambled sequence of instruction would be most detrimental to the learning of low ability students, and that the high ability students could learn effectively by either method. The conclusions drawn from the three investigations were as follows:

1) As predicted, the effects of scrambled sequence were nonsignificant in the case of the anatomy program, but contrary to expectations the detrimental effects of the scrambled sequence were much smaller than had been expected for the hierarchical modern mathematics program. The scrambled sequence increased errors and instructional time in the modern mathematics program (i.e., within program performance measures) but it did not detrimentally affect the level of achievement at the completion of instruction as measured by an immediate posttest. The increase in errors and instructional time within the modern mathematics program supported the contention that the modern mathematics program contained a conceptual hierarchy. An analysis of the error data indicated that students in the scrambled sequence group were able to overcome the effects of the sequence and eventually achieved (apparently through their own organizational strategies) a level of criterion performance comparable to that of the ordered sequence group.

2) An examination of the aptitude by sequence interactions in the experiment employing the modern mathematics program provided tentative support for the hypothesis that a carefully organized sequence of instruction is more important to the learning of low aptitude students than high aptitude students. These results were consistent with the findings of Stolurow (1964) who found that a "mixed" sequence of instruction was more detrimental to the learning of low ability students than high ability students.

References

Florida State University studies computer's teaching potential Data Processing for Education, Dec. 1964, 3, 14-15.

Bitzer, D. L. and Easley, Jr., J. A. Plato: A computer-controlled teaching system. Paper presented DAVI Convention, Rochester, New York April 22, 1964.

Boneau, C. A. The effects of violations of assumptions underlying the t-test. Psychology Bulletin, 1960, 57, 49-64.

Braunfeld, P. G. Problems and prospects of teaching with a computer. J. of Ed. Psy., 1964, 55, 201-211.

Brown, Bobby R. An Instrument for the Measurement of Expressed Attitude Toward Computer Assisted Instruction. In Semi Annual Progress Report, Experimentation with Computer-Assisted Instruction in Technical Education. Project No. 5-85-074, Prepared by Harold E. Mitzel et al., December 31, 1966.

Cronbach, L. J. The two disciplines of scientific psychology. Amer. Psychologist, 1957, 12, 671-684.

Evans, J. L. On investigation of teaching machine variables using learning programs in symbolic logic. Unpublished doctoral dissertation, University of Pittsburgh, 1960.

Gagne, R. M. The acquisition of knowledge. Psychol. Rev., 1962, 69, 355-365.

Glaser, R. Principles of programming. In J. P. Lyrought (Ed.), Programmed Learning. Ann Arbor, Mich. Foundation for Research on Human Behavior, 1961, 7-20.

Hamilton, Nancy R. Effects of logical versus random sequencing of items in an autoinstructional program under two conditions of overt response. J. Educ. Psychol. 1964, 55, 258-266.

Levin, G. R. & Baker, B. L. Item Scrambling in a self-instructional program. J. Educ. Psychol., 1963, 54, 138-143.

Lumsdaine, A. A. Instruments and media of instruction. In N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago, Ill.: Rand-McNally, 1963, 583-682.

Mager, R. F. and Clark, C. Explorations in student-controlled instruction. Psy. Reports, 1963, 13, 71-76.

Maher, Anna. Computer-based instruction (CBI): Introduction to the IBM project. IBM Research Report. RC1114, 1964.

Mitzel, H. E. The development and presentation of four college courses by computer teleprocessing. Interim Report, United States Office of Health, Education, and Welfare, Project #OE-4-16-010. Revised April, 1966.

Mitzel, H. E. and Wodtke, K. H. The development of four different college courses for presentation by computer teleprocessing. Interim Report, United States Office of Education, Department of Health, Education, and Welfare, Project No. OE-4-16-010, June, 1965.

Norton, D. W. An empirical investigation of some effects of non-normality and heterogeneity on the F-distribution. Unpublished doctoral dissertation, State University of Iowa, 1952.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. The Measurement of Meaning. Urbana: University of Illinois Press, 1957.

Payne, David A., Krathwohl, David R., and Gordon, John. The effect of sequence on programmed instruction. Amer. Educational Research Journal, 1967, 4, 125-132.

Roe, K. V. Case, H. W., and Roe, A. Scrambled versus ordered sequence in autoinstructional programs. J. Educ. Psychol., 1962, 53, 101-104.

Skinner, B. F. Teaching machines. Science, 1958, 128, 969-977.

Stolurow, L. W. Social Impact of Programmed Instruction: Aptitudes and Abilities Revisited In J. P. DeCecco (Ed.) Educational Technology, New York: Holt, Rinehart, and Winston, 1964, 348-355.

Stolurow, L. M. and Davis, D. J. Teaching machines and computer-based systems. Urbana, Illinois: University of Illinois, Training Research Laboratory. Technical Report No. 1. Office of Naval Research, NONR 3985 (04), 1963.

Suppes, P. Modern learning theory and the elementary school curriculum. Amer. Ed. Res. J., 1964, 1, 79-93.

Uttal, W. R. My teacher has three arms!!! Yorktown Heights, N. Y.: IBM Research Report RC-788, 1962.

Wodtke, Kenneth H. Scrambled versus ordered course sequencing in computer-assisted instruction: Experiment I In Semi-Annual Progress Report, <u>Experimentation with Computer-Assisted Instruction in Technical Education</u>. Project No. 5-85-074, Prepared by Harold E. Mitzel *et al*., December 31, 1965.

Wodtke, Kenneth H. On the assessment of retention effects in educational experiments. *Journal of Experimental Education*, Summer 1967, *35*, 28-36.

APPENDIX A

A.1 Sample of the Original Number Conversion Program

A.2 Sample of the Revised Number Conversion Program

Appendix A.1

Look for the highest power of eight represented in 724 (ten).
What is it?
2
Eight to the 3rd power is the highest power of 8 in 724 (ten).
Type eight to the third power.
8 3rd power
Eight to the 3rd power is the highest power of 8 in 724 (ten).
Type eight to the third power.
eight to the third power
Correct.

What is the numerical value of eight to the third power?
512
512 is correct.

We subtract 512 from 724, resulting in 212. What power of eight do we look for in this number?
5
We look for eight to the second power. Type eight to the second power.
8 to the second power
Correct.

How many groups of eight to the second power are represented in 212?
1
There are 3 groups of 8 squared in 212. Type 3.
3
Correct.

We subtract 3 x 8 squared, or 192, from 212.
How many groups of 8 to the first power does the result, 20, contain?
2
Correct.

Subtract; the remainder is 4. Therefore, putting all of this work together, what is 724 (ten) in base eight?
3324
724 (ten) = 1324 (eight). Type 1324 (eight).
1324
Correct - 724 (ten) = 1324 (eight).

Look again at the chart in display 5-E. What is the value in base ten of the 211 (eight) shown?
137
Correct

What is the value of 47 (eight) in base ten?
39
Correct

What is the value of 526 (eight) in base ten?
34
Think carefully--it's 5 times 8 squared (the base squared) plus 2 times 8 (the base) plus 6 ones. Now--what's the value of 526 (eight) in base ten?
342
Correct

What is the value of 323 (eight) in base ten?
i don't know
Let's look again. We have 3 groups of sixty-four or 192 plus 2 groups of 8 or 16 plus 3 ones. Total 192 + 16 + 3. What is 323 (eight) in base ten?
361
Let's look again. We have 3 groups of sixty-four or 192 plus 2 groups of 8 or 16 plus 3 ones. Total 192 + 16 + 3. What is 323 (eight) in base ten?
261
Let's look again. We have 3 groups of sixty-four or 192 plus 2 groups of 8 or 16 plus 3 ones. Total 192 + 16 + 3. What is 323 (eight) in base ten?
211
Correct

What is the value of 2372 (eight) in base ten?
gfd
Check your work. 2372 (eight) means 2 groups of 8 to the third power plus 3 groups of 8 to the second power plus 7 groups of 8 to the first power plus 2 ones. That's 2 x 512 plus 3 x 64 plus 7 x 8 plus 2. When you add these together, what is the answer?
1274
Right!

5-30

Let's check back and review a bit. What number in base eight means 5 of the base and 3 ones?
53
Correct - 53 (eight)

What number follows 37 in base eight?
38
Remember the numeral 8 is not used in base eight. Try again.
40
Correct - 40 (eight)

In 652 (eight), which digit shows the number of groups of 8 squared, or eight to the second power.
5
Remember the place value in base eight. Consider 652 (eight). From right to left, 2 is in the ones place or units place. 5 is in the 8 to the first power place. What digit is, therefore, in the 8 squared place?
*6*Right!

What is a logical first step in changing 95 (ten) to base eight? Type a, b, c, or d.
a. Look for the largest power of 8 contained in 95 (ten).
b. Divide by 8.
c. Write "5" in the units column.
d. Change the 9 since there can be no 9 in base eight.
b
This method, involving repeated division by 8, can be meaningful for ascertaining the powers of 8.

Actually, either a or b above were accepted as correct answers. "a" is in accord with the method taught in this program. For a more detailed explanation of the short-cut method in "b", turn to display 5-F. If you decide to skip display 5-F, type go on. Otherwise follow the directions on the display.

Appendix A.2

BEFORE BEGINNING THIS LESSON, READ THE INTRODUCTION IN THE BOOK OF EXHIBITS GIVEN TO YOU. WHEN YOU ARE READY TO PROCEED, PRESS EOB.

1. FOR THE NEXT FEW ITEMS, YOU WILL BE GIVEN A SAMPLE PROBLEM, AND A SAMPLE SOLUTION. BELOW THIS SAMPLE SOLUTION IS A PROBLEM WHICH YOU ARE TO SOLVE.

   SAMPLE PROBLEM: $2(10^2)$ = (SEE SOLUTION BELOW)

   SAMPLE SOLUTION: $2(10^2) = 2(10 \times 10) = 2(100) = 200$

   PROBLEM: $5(10^2) = 5(10 \times 10) = 5(100) =$?

   NOTE: WHENEVER YOU SEE A BLANK SPACE LIKE THIS ? YOU ARE TO TYPE THE ANSWER THAT BELONGS IN THE SPACE, AND THEN PRESS EOB TO ENTER THE ANSWER INTO THE COMPUTER. DO THAT NOW. TYPE THE ANSWER, AND THEN EOB.

5((

NO. THE CORRECT SOLUTION IS:

$5(100) = 500$. 5(100) IS THE SAME AS 5×100.

YOU SHOULD HAVE TYPED: 500.

2. SAMPLE PROBLEM: $3(10^3)$ = (SEE SOLUTION BELOW)

   SAMPLE SOLUTION: $3(10^3) = 3(10 \times 10 \times 10) = 3(1000) = 3000$

   PROBLEM: $4(10^3) = 4(10 \times 10 \times 10) = 4(1000) =$?

   REMEMBER, EVERY TIME YOU SEE A BLANK LIKE THE ONE ABOVE, YOU ARE TO TYPE THE ANSWER THAT BELONGS IN THE SPACE AND THEN PRESS EOB.

4000
VERY GOOD.

3. SAMPLE PROBLEM: $4(10^2) + 3(10^1)$ = (SEE SOLUTION BELOW)

   SAMPLE SOLUTION: $4(10^2) + 3(10^1) = 4(100) + 3(10) = 400 + 30 = 430$

   PROBLEM: $6(10^2) + 2(10^1) = 6(100) + 2(10) = 600 + 20 +$?

620
CORRECT

4. SAMPLE PROBLEM: $2(10^2) + 1(10^1) + 2(10^0) =$ (<u>SEE SOLUTION BELOW</u>)

SAMPLE SOLUTION: $2(10^2) + 1(10^1) + 2(10^0) = 200 + 10 + 2 = 212$

NOTE: 10^0 ALWAYS EQUALS 1; HENCE, $2(10^0) = 2(1)$

PROBLEM: $1(10^2) + 2(10^1) + 3(10^0) =$ ___?___

DO YOUR COMPUTATIONS WITH PENCIL AND PAPER. THEN TYPE THE ANSWER ONLY.

123
CORRECT.

5. IN AN EXPRESSION SUCH AS THIS, $3(10^4)$, THE 3 IS CALLED A COEFFICIENT; THE 10 IS CALLED A BASE, AND THE 4 IS CALLED AN EXPONENT. WE WILL BE USING THESE TERMS IN THIS LESSON; THEREFORE, YOU WILL NEED TO LEARN THEM. FOR PRACTICE, TYPE THE NAME OF THE 3, 10, AND THE 4 BELOW.
COEFFICIENT, BASE, EXPONENT
CORRECT

6. IN THIS PROBLEM, $4(13^5)$, "4" IS CALLED THE _____. "13" IS THE _____. "5" IS THE _____.
COEFFICIENT, BASE, EXPONENT
CORRECT.

7. THE DIGITS WE SEE IN A NUMBER ARE REALLY COEFFICIENTS. (FOR EXAMPLE, IN THE NUMBER "23", THE "2" AND "3" ARE COEFFICIENTS.) THE BASES AND EXPONENTS OF THE COEFFICIENTS ARE USUALLY UNDERSTOOD AND ARE LEFT OUT <u>UNLESS THE NUMBER IS EXPANDED</u>.

SAMPLE PROBLEM: EXPAND 49

SAMPLE SOLUTION: $49 = 4(10^1) + 9(10^0)$

PROBLEM: $81 = 8(10^1) +$ ___?___ (10^0)

YOU WILL NEED TO PROVIDE THE MISSING COEFFICIENT ONLY

1
CORRECT

8. SAMPLE PROBLEM: EXPAND 321

PAY PARTICULAR ATTENTION TO WHAT HAPPENS TO THE COEFFICIENTS 3, 2 AND 1 IN THE SAMPLE SOLUTION.

SAMPLE SOLUTION: $321 = 3(10^2) + 2(10^1) + 1(10^0)$

PROBLEM: $233 = \underline{\ ?\ }\ (10^2) + \underline{\ ?\ }\ (10^1) + \underline{\ ?\ }\ (10^0)$

YOU NEED TO PROVIDE THE MISSING COEFFICIENT.

2, 3, 3,
CORRECT

9. WE HAVE SAID THAT THE DIGITS IN A NUMBER ARE COEFFICIENTS, AND THAT THE BASES AND EXPONENTS ARE UNDERSTOOD. IN THE DECIMAL SYSTEM IT IS UNDERSTOOD THAT THE BASE IS ALWAYS __?__.

10
CORRECT

10. SAMPLE PROBLEM: EXPAND 320

SAMPLE SOLUTION: $320 = 3(10^2) + 2(10^1) + 0(10^0)$
THE NUMBERS 3, 2, and 0 = COEFFICIENTS
THE NUMBERS 10, 10, and 10 = BASE

PROBLEM: $912 = 9(10^2) + 1(10^1) + \underline{\ ?\ }(\underline{\ ?\ })$

YOU NEED TO PROVIDE THE MISSING COEFFICIENT AND THE MISSING BASE. MAKE SURE TO SEPARATE EACH ANSWER WITH THE SPACE BAR.
2 10^0
CORRECT.

APPENDIX B

B.1 Criterion Test for the Number Conversion Programs

B.2 Transfer Test for the Revised Number Conversion Program

B.3 Student Reaction Inventory

B.4 Sample of the Anatomy of the Ear Program

Appendix B.1

(PART I B)

CONO
Date________________

Name________
CAI Student Number________

1. What is 724 (ten) in base eight?

2. What is 2372 (eight) in base ten?

3. What is 321 (five) in base ten?

4. What is 79 (ten) in base five?

5. What is 23 (eleven) in base ten?

6. What is 36 (nine) in base ten?

7. What is 269 (ten) in base five?

8. What is 3 (ten) in base two?

9. What is 10 (ten) in base two?

10. What is 22 (ten) in base two?

11. What does 364 (seven) equal in base 10?

12. What does 324 (eight) equal in base six?

13. What does 115 (fifteen) equal in base ten

14. What does 2 (six) equal in base seven?

15. What does 100 (four) equal in base two?

16. What does 225 (seven) equal in base nine?

17. How is one of the base always written regardless of the value of the base?

18. How is one of the base squared always written?

19. How is one of the base cubed always written?

20. What is the meaning of the 6 in 632 (base nine)?

a. $6(9^1)$

b. $6^2(9^1)$

c. $6(9^2)$

d. $6(9^3)$

e. $6^3(9^1)$

21. What is the meaning of the 5 in 15,342 (base six)?

a. $5(6^1)$

b. $5^4(6^1)$

c. $5^3(6^1)$

d. $5(6^4)$

e. $5(6^3)$

22. What is the expanded form of 3726 (base nine)?

a. $3(9^4) + 7(9^3) + 2(9^2) + 6(9^1)$

b. $9(3^3) + 9(7^2) + 9(2^1) + 6$

c. $3(10^3) + 7(10^2) + 2(10^1) + 6(1)$

d. $3(9^3) + 7(9^2) + 2(9^1) + 6(9^0)$

e. $10(3^3) + 10(7^2) + 10(2^1) + 10(9^0)$

Appendix B.2

POSTTEST (FORM T)

NAME______________________ DATE__________________

Most people think that a number system must necessarily employ numerals; however, if the basic principles of number systems are understood any symbols may be used. For example, the hypothetical "fum society" uses the following number names with the base equal to fum:

		x	x	x x
none	x	x	x x	x x
fizzle	fe	fi	fo	fum

The base squared = gum

base cubed = hum

base to the fourth = jum.

1. Write the number in fum notation which is equivalent to the numeral 1233 (in base four). Use commas to separate the place value positions.

2. Write the numeral 1312 (base four) in fum notation:

3. Show how we would count twenty objects in the fum system.

4. Add the following numbers in the fum system. <u>Do not</u> convert to base ten, but add the two numbers directly. Show all your work.

	fe-gum, fi-fum, fo
+	, fo-fum, fi

a. What value (in <u>fum</u> notation) is carried in column one? ____________

b. What value is carried in column two? ____________

5. Try adding:

	fo-gum, fe-fum, fe
+	fi-gum, fi-fum, fo

a. Carried over from column one (in fum notation) = ____________
b. Carried over from column two (in fum notation) = ____________
c. Carried over from column three (in fum notation) = ____________

6. Let's try adding some numerical values in the base eight:

	4 7 2 (eight)
+	6 5 4 (eight)

a. Value carried from column one = ____________
b. Value carried from column two = ____________
c. Value carried from column three = ____________

7. Subtract:

$$\begin{array}{rl} 3\ 2\ 5 & \text{(base eight)} \\ -\ \underline{2\ 7\ 6} & \underline{\text{(base eight)}} \end{array}$$

a. Value borrowed in column one ______________________________

b. Value borrowed in column two ______________________________

8. Add the following numbers in base <u>two</u>.

$$\begin{array}{rl} 1\ 1\ 0\ 1 & \text{(two)} \\ +\ \underline{1\ 0\ 1\ 1} & \text{(two)} \end{array}$$

a. Carried from column one ______________________________

b. Carried from column two ______________________________

c. Carried from colu..n three ______________________________

$$\begin{array}{rl} 1\ 0\ 0\ 1 & \text{(two)} \\ +\ \underline{\ \ 1\ 1\ 1} & \text{(two)} \end{array}$$

a. Carried from column one ______________________________

b. Carried from column two ______________________________

c. Carried from column three ______________________________

8. (continued)

$$\begin{array}{r} 1\ 0\ 1\ 0\ 1 \text{ (two)} \\ +\ \underline{1\ 0\ 1\ 1} \text{ (two)} \end{array}$$

a. Carried from column one ________________

b. Carried from column two ________________

c. Carried from column three ________________

d. Carried from column four ________________

$$\begin{array}{r} 1\ 1\ 1\ 1 \text{ (two)} \\ -\ \underline{1\ 0\ 1} \text{ (two)} \end{array}$$

a. Borrowed from column one ________________

b. Borrowed from column two ________________

c. Borrowed from column three ________________

$$\begin{array}{r} 1\ 0\ 0\ 1 \text{ (two)} \\ -\ \underline{1\ 1\ 1} \text{ (two)} \end{array}$$

a. Borrowed from column one ________________

b. Borrowed from column two ________________

c. Borrowed from column three ________________

8. (continued)

$$\begin{array}{rl} 1\ 0\ 1\ 1 & \text{(two)} \\ -\ \underline{\ \ 1\ 0\ 1} & \text{(two)} \end{array}$$

a. Borrowed from column one ____________

b. Borrowed from column two ____________

c. Borrowed from column three ____________

9. Add the following <u>base</u> <u>five</u> numbers:

$$\begin{array}{r} 3\ 4\ 2 \\ +\ \underline{1\ 1\ 4} \end{array}$$

carried in column one ____________

carried in column two ____________

carried in column three ____________

$$\begin{array}{r} 2\ 4\ 0\ 1 \\ +\ \underline{1\ 3\ 4\ 2} \end{array}$$

carried in column one ____________

carried in column two ____________

carried in column three ____________

carried in column four ____________

$$\begin{array}{r} 3\ 2\ 4\ 3 \\ +\ \underline{4\ 3\ 3\ 2} \end{array}$$

carried in column one ____________

carried in column two ____________

carried in column three ____________

carried in column four ____________

10. Count to fifteen in the binary system.

Appendix B.3

STUDENT ATTITUDE TOWARD COMPUTER ASSISTED INSTRUCTION

This is not a test of information; therefore, there is no one "right" answer to a question. We are interested in your opinion on each of the statements below. Your opinions will be strictly confidential. Do not hesitate to put down exactly how you feel about each item.

NAME______________________________ DATE____________________

CIRCLE THE RESPONSE THAT MOST NEARLY REPRESENTS YOUR REACTION TO EACH OF THE STATEMENTS BELOW.

1. The method by which I was told whether I had given a right or wrong answer became montonous.

   Strongly Disagree — Disagree — Uncertain — Agree — Strongly Agree

2. The material presented to me by Computer Assisted Instruction caused me to feel that no one really cared whether I learned or not.

   Strongly Disagree — Disagree — Uncertain — Agree — Strongly Agree

3. While taking Computer Assisted Instruction I felt challenged to do my best work.

   Strongly Disagree — Disagree — Uncertain — Agree — Strongly Agree

4. While taking Computer Assisted Instruction I felt isolated and alone.

   All the time — Most of the time — Some of the time — Only occasionally — Never

5. While taking Computer Assisted Instruction I felt as if someone were engaged in conversation with me.

   All the time — Most of the time — Some of the time — Only occasionally — Never

6. As a result of having studied some material by Computer Assisted Instruction, I am interested in trying to find out more about the subject matter.

   Strongly Disagree — Disagree — Uncertain — Agree — Strongly Agree

7. I was more involved in running the machine than in understanding the material.

All the time | Most of the time | Some of the time | Only occasionally | Never

8. Computer Assisted Instruction makes the learning too mechanical.

Strongly Disagree | Disagree | Uncertain | Agree | Strongly Agree

9. I felt as if I had a private tutor while on Computer Assisted Instruction.

Strongly Disagree | Disagree | Uncertain | Agree | Strongly Agree

10. I found it difficult to concentrate on the course material because of the hardware.

All the time | Most of the time | Some of the time | Only occasionally | Never

11. The Computer Assisted Instruction situation made me feel quite tense.

Strongly Disagree | Disagree | Uncertain | Agree | Strongly Agree

12. Computer Assisted Instruction is an inefficient use of the student's time.

Strongly Disagree | Disagree | Uncertain | Agree | Strongly Agree

13. Concerning the course material I took by Computer Assisted Instruction, my feeling toward the material before I came to Computer Assisted Instruction, was

Very Favorable | Favorable | Indifferent | Unfavorable | Very unfavorable

14. Concerning the course material I took by Computer Assisted Instruction, my feeling toward the material after I had been on Computer Assisted Instruction.

Very Favorable | Favorable | Indifferent | Unfavorable | Very unfavorable

15. While on Computer Assisted Instruction I encountered mechanical malfunctions.

Very often | Often | Occasionally | Seldom | Very Seldom

16. I felt frustrated by the Computer Assisted Instruction situation.

Strongly Disagree | Disagree | Uncertain | Agree | Strongly Agree

17. The Computer Assisted Instruction approach is inflexible.

Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree

18. Even otherwise interesting material would be boring when presented by Computer Assisted Instruction.

Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree

19. In view of the effort I put into it, I was satisfied with what I learned while taking Computer Assisted Instruction.

Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree

20. In view of the amount I learned, I would say Computer Assisted Instruction is superior to traditional instruction.

Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree

21. With a course such as I took by Computer Assisted Instruction, I would prefer Computer Assisted Instruction to traditional instruction.

Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree

22. I am not in favor of Computer Assisted Instruction because it is just another step toward de-personalized instruction.

Strongly Disagree	Disagree	Uncertain	Agree	Strongly Agree

8/18/67
Form 1

Appendix B.4

a-1. Following the normal pathway of sound, type the names of the three main divisions of the ear.
outer ear, inner ear, and middle ear
Good. Another name for outer ear is external ear.

a-2. What is the name for that part of the external ear most visible and on the outside of the head.
pinna
Correct. The terms pinna and auricle are synonymous.

a-3. What small structure of the ear divides the external ear from the middle ear. You need a speculum to see it.
ear drum
Correct. However a more technical term we shall be using is tympanic membrane.

a-4. What passageway connects the pinna or auricle with the tympanic membrane.
auditory canal
This is one of several correct answers, however, the preferred term is external auditory meatus.

a-5. What is the name of the substance secreted by the glands along the external auditory meatus?
cerutin
Cerumen is correct.

We have briefly glanced at the anatomy of the outer ear. Now let's consider something about its function. Answer the following multiple choice questions by typing the letter of the correct choice.

a-6. The function of the pinna or auricle in human beings can be described as:
a. Sound concentration
b. Support for women's earrings
c. Sound directing
d. Basically useless
d.
Type the lower case letter of the correct choice.
d
Anyone of the four alternatives has some truth in it. However, look at the fifth line from the top of page 18 in Newby for an important consideration in respect to this question.

a-7. On humans the external auditory meatus has which of the following main functions for hearing?

a. A receptacle for cerumen
b. Aerate the tympanic membrane
c. Equalization of pressure on both sides of the t.m.
d. Allows sound to arrive at the t.m. for further hearing functions

b
Though this, in fact, may occur, it is not the primary function of the external auditory meatus. Try again.

c
Though this, in fact, may occur, it is not the primary function of the external auditory meatus. Try again.

a
Though this, in fact, may occur, it is not the primary function of the external auditory meatus. Try again.

d
Correct

a-8. What does one normally find along the external auditory meatus other than cerumen?
hair
Correct.

a-9. What is the name of part O of the model and your handout? A helical spring is a coiled spring. The part of
furl
Feel the outer and upper edge of your auricle. Follow it around. Find the comparable part in Newby Figure 2-3 and type its name.
helix
Good. Notice that the helix extends around a portion of the ear--hence there are four letters O on the model. Write "helix" on the line next to O on your handout.

a-10. Identify part M on your handout. This bump on the helix is not well developed on the model, but the M is there. You may be able to find a bump like this on your own helix. Charles Darwin believed that this tubercle was a holdover in man from an earlier form of his development. Type the name of part M.
tubercle
Reread the question. The answer is evident in the question.
bump
The answer is Darwin's tubercle. Type it.
Darwin's tubercle
The correct answer is Darwin's tubercle. Write it on the line next to M on your handout.

a-11. Follow your own helix up, around, and to the front until it attaches to your head. This is part E on your handout. What do we call it?

?

Incorrect. Keep trying.

helix curl

Type: crest of helix

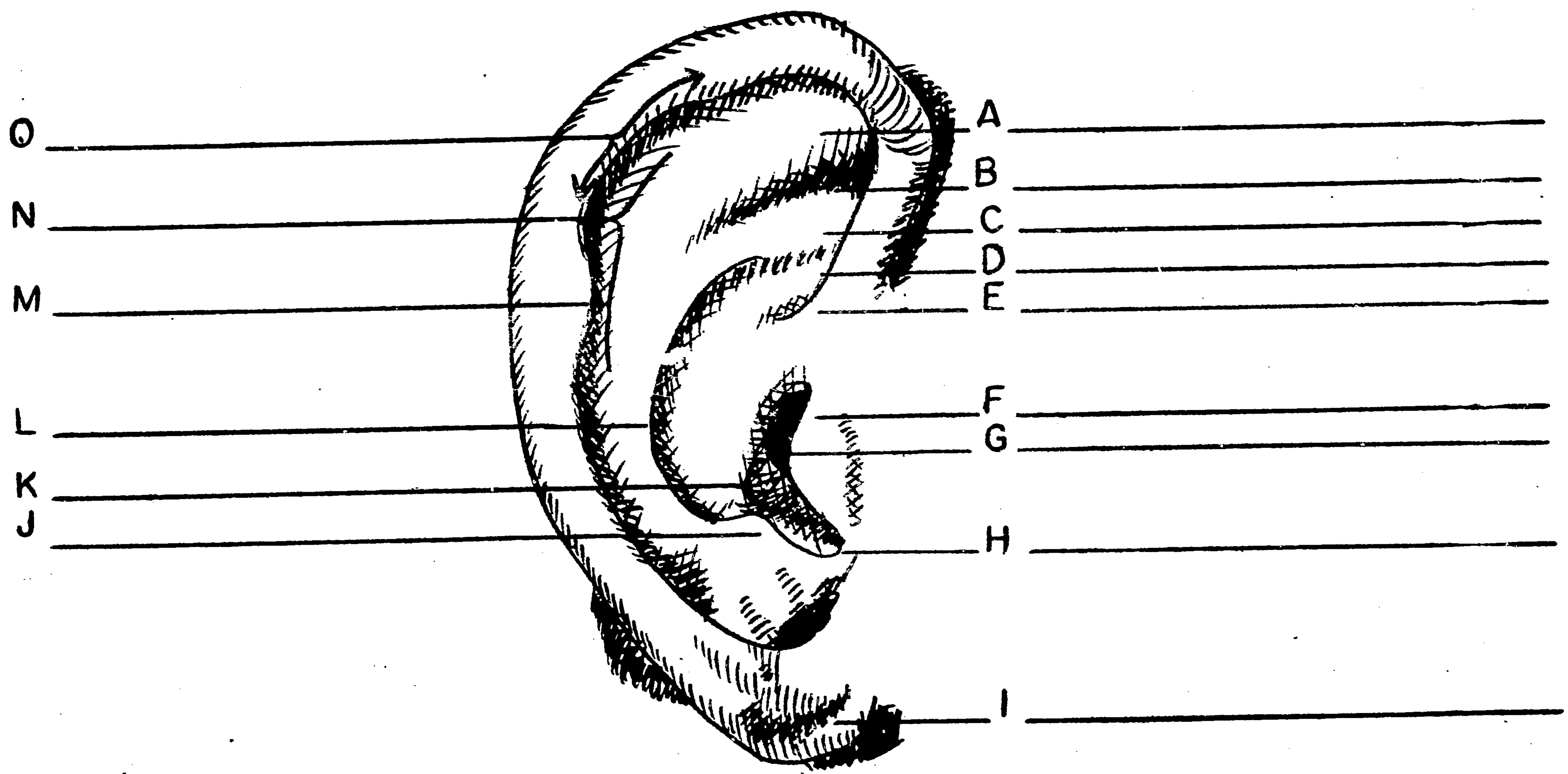
crest of helix

Correct. This anatomical part has at least three names, crest of helix, limb of helix, or crus of helix. Label part E on your handout.

Appendix C

C.1 Posttest for the Anatomy of the Ear Program

THE RIGHT HUMAN PINNA



CAI SPA 430
10/9/65

Appendix C.1

Date________________________ Name:______________________________

THE EAR

1. ______________ What is another technical anatomical name for the pinna?

2. ______________ What is another name for the external ear canal? (three words)

3. ______________ What is the technical name for earwax?

4. ______________ In humans, the external auditory meatus has which of the following main functions for hearing?

    a. a receptacle for cerumen

    b. aerate the tympanic membrane

    c. equalization of pressure on both sides of the t.m.

    d. allows sound to arrive at the t.m. for further hearing function

*Note: Ss' task was to fill in the blanks on this form and to label the picture of the ear on the next page.